

This listing of claims will replace all prior versions, and listings, of claims in the Application:

**LISTING OF CLAIMS:**

Claim 1 (currently amended): A method for determining a genotype associated with increased ~~or decreased~~ resistance to familial bipolar affective disorder in a family affected by bipolar affective disorder, comprising:

determining the genotype of at least one family member, wherein the genotype is determined with at least one marker for at least one chromosomal region linked to a locus associated with resistance to bipolar affective disorder, wherein the chromosomal regions are inclusive of and localized between *D4S402* and *D4S424*; inclusive of and localized between *D4S431* and *D4S404*; or inclusive and localized between *D11S394* and *D11S29*;

determining, after the age of onset, the bipolar affective disorder disease status in the family member;

comparing the genotype with the bipolar affective disorder disease status; and  
determining therefrom the genotype associated with increased ~~or decreased~~ resistance to bipolar affective disorder.

Claim 2 (original): The method of claim 1, wherein the genotype is determined with markers for at least two of the chromosomal regions.

Claim 3 (original): The method of claim 2, wherein the genotype is determined with markers for three of the chromosomal regions.

Claim 4 (original): The method of claim 1, wherein the chromosomal region is inclusive of and localized between markers *D4S422* and *D4S1625*.

Claim 5 (original): The method of claim 4, wherein the marker is *D4S175*, *D4S422*, *D4S1576*, *D4S2294*, *D4S1579*, *D4S397*, *D4S3089*, *D4S2965*, *D4S192*, *D4S420*, *D4S1644*, *D4S3334*, or combinations thereof.

Claim 6 (original): The method of claim 1, wherein the chromosomal region is inclusive of and localized between markers *D4S3007* and *D4S419*.

Claim 7 (original): The method of claim 6, wherein the marker is *D4S3007*, *D4S394*, *D4S2983*, *D4S2923*, *D4S615*, *AFM<sub>q</sub>184za9*, *D4S2928*, *D4S1065*, *D4S1582*, *D4S107*, *D4S3009*, *D4S2906*, *D4S2949*, *AFM087zg5*, *D4S2944*, *D4S403*, *D4S2942*, *D4S2984*, *D4S1602*, *D4S1511*, *D4S2311*, *D4S3048*, or combinations thereof.

Claim 8 (original): The method of claim 7, wherein the marker is *D4S3009*, *D4S2906*, *D4S2949*, *AFM087zg5*, *D4S2944*, *D4S403*, *D4S2942*, *D4S2984*, *D4S1602*, *D4S1511*, *D4S2311*, or combinations thereof.

Claim 9 (original): The method of claim 1, wherein the chromosomal region is inclusive of and localized between markers *D11S133* and *D11S29*.

Claim 10 (original): The method of claim 9, wherein the marker is *D11S133*, *D11S147*, *CD3D*, *D11S285*, *D11S29*, or combinations thereof.

Claim 11 (original): The method of claim 1, wherein the genotype at a single chromosomal region is determined with at least three markers.

Claim 12 (original): The method of claim 1, wherein the marker is for a restriction fragment length polymorphism or microsatellite polymorphism.

Claim 13 (withdrawn)

Claim 14 (withdrawn)

Claim 15 (original): The method of claim 1, wherein the marker is amplified.

Claim 16 (original): The method of claim 15, wherein the marker is amplified by the polymerase chain reaction.

Claim 17 (original): The method of claim 1, wherein the presence or absence of an allele associated with increased resistance to bipolar affective disorder is determined.

Claim 18 (original): The method of claim 1, wherein the genotype of an affected family member is determined.

Claim 19 (original): The method of claim 1, wherein the genotype of a non-affected family member is determined.

Claim 20 (original): The method of claim 1, further comprising:

determining the genotype of at least one family member, wherein the genotype is determined with at least one marker for at least one chromosomal region linked to a locus associated with susceptibility to bipolar affective disorder, wherein the chromosomal regions are inclusive of and localized between *D6S344* and *D6S89*; inclusive of and localized between *D13S171* and *D13S218*; or at about *D15S148*.

Claim 21 (currently amended): The method of claim 1, further comprising:

determining the genotype of a tested individual from the affected family, wherein the genotype is determined with at least one marker for at least one chromosomal region linked to a locus associated with resistance to bipolar affective disorder, wherein the chromosomal regions are inclusive of and localized between *D4S402* and *D4S424*; inclusive of and localized between *D4S431* and *D4S404*; or inclusive and localized between *D11S133* and *D11S29*;

comparing the genotype of the tested individual to the genotype associated with increased ~~or decreased~~ resistance to bipolar affective disorder; and

determining therefrom the ~~increased or decreased~~ risk of the tested individual developing familial bipolar affective disorder.

Claim 22 (original): The method of claim 21, wherein the genotype of the tested individual is compared to the genotype of an affected family member.

Claim 23 (currently amended): A method for determining the contribution of a chromosomal region to the presence ~~or absence~~ of resistance to bipolar affective disorder in a family affected by bipolar affective disorder, comprising:

determining the corresponding genotype of at least two family members, wherein the genotype is determined with at least one marker for at least one tested chromosomal region linked to a locus associated with resistance to bipolar affective disorder,

wherein the tested chromosomal regions are inclusive of and localized between *D4S402* and *D4S424*; inclusive of and localized between *D4S431* and *D4S404*; or inclusive and localized between *D11S133* and *D11S29*;

determining, after the age of onset, the bipolar affective disorder disease status in the family members;

comparing the genotypes of the family members; and

determining therefrom the contribution of the chromosomal region to the presence ~~or absence~~ of resistance to bipolar affective disorder in the family.

Claim 24 (currently amended): A method for determining a genotype associated with increased susceptibility ~~or decreased resistance~~ to familial bipolar affective disorder in a family affected by bipolar affective disorder, comprising:

~~determining the genotype of at least one family member, wherein the genotype is determined with at least one marker for at least one chromosomal region linked to a locus associated with resistance to bipolar affective disorder, wherein the chromosomal regions are inclusive of and localized between *D4S402* and *D4S424*; inclusive of and localized between *D4S431* and *D4S404*; or inclusive and localized between *D11S133* and *D11S29*;~~

determining the genotype of at least one family member, wherein the genotype is determined with at least one marker for at least one chromosomal region linked to a locus associated with susceptibility to bipolar affective disorder, wherein the chromosomal regions are inclusive of and localized between *D6S344* and *D6S89*; inclusive of and localized between *D13S171* and *D13S218*; or at about *D15S148*;

determining, after the age of onset, the bipolar affective disorder disease status in the family member;

comparing the genotype with the bipolar affective disorder disease status; and

determining therefrom the genotype associated with increased susceptibility ~~or decreased resistance~~ to bipolar affective disorder.

Claim 25 (original): The method of claim 24, wherein the marker associated with susceptibility is *D6S7*, *D13S1*, *D15S45*, or combinations thereof.

Claim 26 (currently amended): The method of claim 24, further comprising:

determining the genotype of a tested individual from the affected family,

wherein the genotype is determined with at least one marker for at least one chromosomal region linked to a locus associated with resistance to bipolar affective disorder, wherein the chromosomal regions are inclusive of and localized between *D4S402* and *D4S424*; inclusive of and localized between *D4S431* and *D4S404*; or inclusive and localized between *D11S133* and *D11S29*;

comparing the genotype of the tested individual to the genotype associated with increased susceptibility ~~or decreased resistance~~ to bipolar affective disorder; and

determining therefrom the increased or decreased risk of the tested individual developing familial bipolar affective disorder.

Claim 27 (withdrawn)

## REMARKS/ARGUMENTS

Claims 1-27 are pending. Claims 1, 21, 23, 24 and 26 have been amended. Claims 13, 14 and 27 have been withdrawn.

The substance of claim 1 has been divided between claims 1 and 24. No new matter has been added.

The documents to which applicant wishes to claim priority are properly indicated in the first sentence of the application.

In the specification, the paragraphs beginning at page 22, line 1 and continuing through page 23, line 10, have been amended to specify the SEQ ID NO: corresponding to each sequence. This amendment renders moot the objection with regard to "Sequence Rules." Support for this amendment is found, *e.g.*, in the substitute sequence listing filed June 13, 2001.

### Rejections under 35 USC §112

At the outset, it is noted that the Examiner has apparently misunderstood the claimed invention. The enablement and the written description rejections appear to be directed to claims that recite a genotype that is associated with increased or decreased resistance to bipolar disorder (*e.g.*, the *result* of a method for determining a genotype associated with increased or decreased resistance to familial bipolar affective disorder (BPAD)). In fact, the claims under consideration are directed to methods, *e.g.*, a method for determining a genotype associated with increased resistance to (protection from) familial bipolar affective disorder (BPAD) ... (independent claim 1); a method for determining the contribution of a chromosomal region to the presence of resistance to BPAD... (independent claim 23); or a method for determining a genotype associated with increased susceptibility to familial bipolar affective disorder (BPAD) ... (independent claim 24). The specification provides both enablement and written description for performing such methods.

The Examiner appears to doubt applicant's assertion in the specification that markers identified therein exhibit statistically significant linkage to either susceptibility to BPAD (*e.g.*, *D6S7* on chromosome 6; *D13S1* on chromosome 13; and *D15S45* on chromosome 15) or resistance to BPAD (*e.g.*, *D4S2949* on chromosome arm 4p; *D4S397* on chromosome arm 4q; and *D11S133* and *D11S29* on chromosome 11). However, the Examiner has presented no reasons or evidence to doubt this assertion.

In fact, the specification clearly shows, *e.g.*, that *D6S7*, *D13S1* and *D15S45* are each linked to susceptibility to BPAD with a SIBPAL p value of 0.0003 (see, *e.g.*, Figures 3, 4 and 5, respectively). Furthermore, the specification shows that *D4S2949* is linked to resistance to (protection from) BPAD with a SIBPAL p value of  $< 5 \times 10^{-5}$  (see, *e.g.*, Figure 6 and page 44, lines 11-13); *D4S2949* is linked to resistance to BPAD with a SIBPAL p value of  $< 0.0001$  (see, *e.g.*, Figure 7); and *D11S133* and *D11S29* are linked to resistance to BPAD with a nominal p value of  $< 5 \times 10^{-5}$  (see, *e.g.*, page 44, lines 15-16). One of skill in the art would know that a SIBAL p value (statistical significance of the genetic linkage between markers based on sib-pair analysis) of  $< 0.001$  is statistically significant under these circumstances. See also the specification at page 34, lines 13-15 (“those markers having a test statistic of p  $< 0.001$  in any one analysis type ... showed evidence for linkage”).

Note that p values are the appropriate statistical criteria for a study of common traits, such as BPAD. LOD scores, on the other hand, are an appropriate statistical tool for evaluating rare disorders. As the specification states on page 34, lines 27-31, although the lod scores in the present study did not reach a LOD=3 criterion, “p values of less than 0.001 (and even 0.0001, which is asymptotically equivalent to  $Z_{\max}=3$ ) are found. Together, these results lend further support to the significance of these intervals as candidate regions.” See also the attached articles by Lander and Schork (Lander *et al.* (1994) *Science* 265, 2037-2038) and Lander and Kruglyak (Lander *et al.* (1995) *Nature Genetics* 11, 241-247) (especially page 244), both of which were published before the filing date of the present application, for discussions of the two statistical methods, and when it is appropriate to use each.

The present specification is replete with sophisticated analyses of the data, employing a variety of statistical techniques, which support the significance of the linkages. The allegation that earlier studies of linkage in BPAD may have been faulty (*e.g.*, as reported in the Berrettini reference cited by the Examiner, and as acknowledged in the specification, for example at page 2, line 27 to page 3, line 2) does not cast doubt on the accuracy of the present findings.

In the methods recited in the instant claims, family members are tested for the presence of markers that flank the markers discussed above, *e.g.*, that lie about 10 cM on either side of one of those markers. See, *e.g.*, the specification at page 14, lines 29-31 and at page 15, lines 5-7. For example, Claim 1 recites determining the genotype with at least one marker for a region inclusive and localized between *D4S402* and *DS424* (on chromosome arm 4q); or inclusive and localized between *D4S431* and *DS404* (on chromosome arm 4p). As

shown in Fig. 7 (for chromosome 4q), marker *D4S402* lies ~32 cM from the marker noted above, *D4S397* (~50 cM - ~18 cM); and, on the other side of this marker, *D4S424* lies ~ 10 cM away (~60 cM - ~50 cM). As shown in Fig. 6 (for chromosome 4p), marker *D4S431* lies ~15 cM from the marker noted above, *D4S2949* (~31 cM - ~15 cM); and, on the other side of this marker, *D4S404* lies ~ 17 cM away (~48 cM - ~31 cM). Shorter spans of markers are recited in dependent claims, as are individual markers within these spans.

One of skill in the art would know that if a particular marker has been shown to be linked to, for example, resistance to (protection from) BPAD, markers lying within a span of about 10 cM on either side of the marker represent likely candidates for further markers exhibiting a linkage to the condition. Moreover, a skilled worker would know that markers even more distant (*e.g.*, at least as far as about 30 cM) would still be considered to be linked to the marker in question and thus would represent a suitable region for additional candidate markers linked to that condition. In the present application, linkage studies indicate that several of the suggested regions do contain markers, in addition to those discussed above, that are linked to susceptibility or resistance to (protection from) BPAD. For example, Fig. 6 shows that at least markers *D4S2366*, *D4S1582*, *D4S419* and *D4S404* on chromosome arm 4p have a SIBPAL p value of < 0.001.

In the method of claim 1, for example, a genotype of a (at least one) family member is determined as discussed above; the BPAD status of the family member is determined; the genotype is compared with the BPAD status; and from this comparison, it is determined if the genotype is associated with increased resistance to BPAD.

The specification is fully enabling for the recited methods: it teaches how to obtain samples from family members (*e.g.*, at page 19, lines 12-27 and Example I); how to perform the genotypic analysis (*e.g.*, in Example II); how to assess BPAD status (*e.g.*, in Examples I and IV) and how to analyze statistically whether the genotype determined is associated with increased or decreased resistance for BPAD (*e.g.*, in Example III). The specification teaches a variety of markers that can be used in the claimed methods, and teaches how additional markers can be generated (*e.g.*, at page 23, line 27 to page 27, line 10). The Office Action states at page 7 that "While one could conduct additional experimentation to determine whether markers exist within the recited regions on chromosome 4 and 11 and these newly discovered markers are associated with BPAD, the outcome of such research cannot be predicted." That is to some extent correct. The methods of the invention are directed in part to determining whether such markers are, in fact, associated with BPAD.



Not only is the specification enabling for the claimed invention, but it also provides written description. Contrary to the allegation of the Examiner on page 10 of the Office Action, the specification provides even more than a “representative number” of markers within each of the claimed regions of the chromosomes. For example, some suitable markers to test for resistance to BPAD, flanking *D4S2949* (on chromosome arm 4p), are indicated at page 14, line 31 to page 15, line 4; some suitable markers flanking *D4S397* (on chromosome 4q) are indicated at page 15, lines 8-9; and some suitable markers on chromosome 11, in the ~20 cM span between *D11S133* and *D11S29*, are indicated at page 15, line 13. See also further candidate markers for resistance indicated on page 18, lines 5-13. Examples of suitable primers for identifying the above markers are indicated on page 22, line 1 to page 23, line 10. The specification also teaches some suitable markers to test for susceptibility for BPAD, and methods for identifying those markers.

As for the allegation in the Office Action on page 7 that “The ability to screen for a wellness allele is even more unpredictable because it is very difficult to distinguish between the presence of a protective allele and the absence of a susceptibility allele,” the specification clearly teaches how to distinguish between these two possibilities. For example, the specification states at page 39, lines 9-16 that

Importantly, because of the long-term, longitudinal nature of the study, even the unaffected, mentally healthy individuals (those without any psychiatric illness) in these families have been closely followed, many for a period of years past the age of risk for BPAD. Consequently, rather than limit this genome-wide search to identifying susceptibility loci for the disease phenotype (BPAD), we tested the hypothesis that “protective” alleles may contribute to the absence of psychiatric illness (*i.e.*, mental health ‘wellness’) in unaffected family members in the ‘high risk’ pedigrees.

and at page 51, lines 13-16:

Accordingly, an important step in our study which demonstrates that there are “protective” alleles was to show that there are “mentally healthy” individuals who share marker alleles that should increase the risk of developing BPAD, and yet, in the presence of ‘protective’ alleles these individuals do not manifest BPAD.

#### The anticipation rejection

The Blackwood reference cited by the Examiner as allegedly anticipating the claimed invention reports that certain markers on chromosome 4p, in the vicinity of marker *D4S394*, are associated with susceptibility to BPAD. By contrast, the present inventors have identified a region close to that identified by Blackwood, which is linked to resistance to (protection

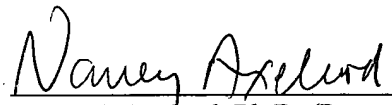
from) BPAD. Applicants believe that the methods recited in the claims are clearly distinguished from Blackwood. Nevertheless, in an effort to expedite prosecution, claims that recite the use of certain markers on chromosome 4p have been amended to clarify that the methods are for determining a genotype associated with increased resistance (rather than increased susceptibility) to BPAD (claim 1) or for determining the contribution of a chromosomal region to the presence of resistance to BPAD (claim 23). Blackwood does not disclose a method for determining a genotype associated with increased resistance to BPAD. In order to anticipate a claim, a reference must disclose all material elements of the claim (*In re Marshall*, 198 USPQ 344 (CCPA 1978)). That is clearly not the case here. (Claim 24, which recites a method for determining a genotype associated with increased *susceptibility* to BPAD does not recite markers in the region studied by Blackwood, and for at least this reason is not anticipated by Blackwood.)

In view of the preceding amendments and arguments, the application is believed to be in condition for allowance, which action is respectfully requested.

Should any additional fee be deemed due, please charge such fee to our Deposit Account No. 22-0261, referencing docket number 31671-186347 and advise us accordingly. However, the Commissioner is not authorized to charge the cost of the issue fee to the Deposit Account.

Respectfully submitted,

Date: July 22, 2003

  
Nancy J. Axelrod, Ph.D. (Patent Agent)  
Registration No. 44,014

VENABLE  
P.O. Box 34385  
Washington, D.C. 20043-9998  
Telephone: (202) 962-4800  
Direct: (202) 962-8334  
Telefax: (202) 962-8300

DC2-DOCS1-471690v2

# Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results

Eric Lander<sup>1,2</sup> & Leonid Kruglyak<sup>1</sup>

Genetic studies are under way for many complex traits, spurred by the recent feasibility of whole genome scans. Clear guidelines for the interpretation of linkage results are needed to avoid a flood of false positive claims. At the same time, an overly cautious approach runs the risk of causing true hints of linkage to be missed. We address this problem by proposing specific standards designed to maintain rigor while also promoting communication.

<sup>1</sup>Whitehead  
Institute for  
Biomedical  
Research, 9  
Cambridge Center,  
Cambridge  
Massachusetts  
02142, USA  
<sup>2</sup>Department of  
Biology,  
Massachusetts  
Institute of  
Technology,  
Cambridge  
Massachusetts  
02139, USA

Genetic dissection of complex traits is becoming central to mammalian genetic analysis. In the fifteen years since it was recognized that genetic inheritance can be traced with naturally occurring DNA sequence variation<sup>1</sup>, the identification of genes responsible for simple mendelian traits has become a straightforward, if still demanding, task. Over 500 such genes have been mapped to specific chromosomal regions in the human and more than 60 have been cloned based on their position. These breakthroughs are steadily reshaping biological and medical thinking. Yet, many of the most important medical conditions — including heart disease, hypertension, diabetes, asthma, schizophrenia, and manic depression — show much murkier inheritance patterns. The geneticists' challenge is now to tease apart the multifactorial causes of these diseases.

In principle, the solution is clear. Genetic mapping of any trait — simple or complex — boils down to finding those chromosomal regions that tend to be shared among affected relatives and tend to differ between affecteds and unaffecteds. Conceptually, this amounts to a three-step recipe: scan the entire genome with a dense collection of genetic markers; calculate an appropriate linkage statistic  $S(x)$  at each position  $x$  along the genome; and identify the regions in which the statistic  $S$  shows a significant deviation from what would be expected under independent assortment.

Yet, these deceptively simple instructions conceal a thorny question: since the statistic  $S(x)$  fluctuates substantially just by chance across an entire genome scan, what constitutes a 'significant' deviation? What standard should be required for declaring linkage?

Although biologists often greet statistical issues with glazed-eyed indifference, we believe that the resolution of this particular question has important consequences

for the future of our field. To reach our goal, geneticists must chart a prudent course between Scylla and Charybdis.

Adopting too lax a standard guarantees a burgeoning literature of false positive linkage claims, each with its own gene symbol (ASTH56, ASTH57, ...). Scientific disciplines erode their credibility when a substantial proportion of claims cannot be replicated — even more so when the claims reach not only the professional journals but also the evening news. Psychiatric genetics provides a cautionary tale, in which a spate of non-replicable findings in the mid-1980s undermined support for such studies<sup>2-7</sup>. It is thus essential that there be a sufficiently stringent standard that linkage is claimed only when there is a high likelihood that the assertion will stand the test of time.

On the other hand, adopting too high a hurdle for reporting results runs the risk that the nascent field will be stillborn. Initial genetic analyses may fall short of the strict threshold for statistical significance, but may nonetheless point to important regions deserving intensive investigation. Without channels by which investigators can report such tentative hints of linkage, the discovery of disease genes may be delayed in an overzealous attempt to avoid all error.

Striking the right balance requires both a mathematical understanding of how often positive results will occur just by chance and a value judgment about the relative costs of false positives and false negatives. Our goal here is to provide an accessible treatment of the first subject and to offer a concrete proposal regarding the second.

## Statistical significance in genome-wide scans

In searching for disease genes, it is important to distinguish between pointwise significance levels and genome-wide significance levels. The pointwise (also called nominal) significance level is the probability that one would encounter such an extreme deviation at a specific locus just by chance. The genome-wide significance level is the probability that one would encounter a deviation somewhere in a whole genome scan. The former concerns a single test of the null hypothesis of no linkage; the latter involves fishing over a large number of tests to find the most significant result.

Consider the following idealized sib pair study. An investigator collects  $n$  pairs of affected sibs, genotypes them using a perfect genetic map that is fully informative at every point in the genome, and calculates the average proportion  $\pi(x)$  of alleles shared identical-by-descent at each location  $x$  in the genome. Geneticists traditionally report the results at each location in one of three essentially equivalent ways: a Z-score, a lod score, or a P value. The Z-score is the number of stan-

## progress

## Box 1 Going to extremes

How often will a linkage study find a significant region? In a whole genome scan, the number of regions tested is large, and the chance of finding a significant region by chance is high. The asymptotic theory described in Box 1 provides a simple formula for the expected number of regions in which a linkage statistic exceeds a threshold  $T$  is given by a simple formula  $\mu(T)$ , explained in Box 1. In fact, the number of such regions approximately follows a Poisson distribution with mean  $\mu(T)$ , and the chance of finding at least one such region is thus  $1 - e^{-\mu(T)} \approx \mu(T)$  when  $\mu(T)$  is small. The approximation becomes asymptotically exact when the number of sib pairs is large and the threshold  $T$  is high. In fact, it is accurate enough for practical purposes provided that  $n$  is at least 50. (It is worth noting that, while lod score analysis of a small number of large families can be quite sensitive to changes in a few key data points, non-parametric statistics based on a large number of small families tend quickly to normal distributions and tend to be robust.)

The mathematical theory of large deviations holds the answer, as was pointed out a few years ago<sup>8,9</sup>. The expected number of chromosomal regions in which a linkage statistic exceeds a threshold  $T$  is given by a simple formula  $\mu(T)$ , explained in Box 1. In fact, the number of such regions approximately follows a Poisson distribution with mean  $\mu(T)$ , and the chance of finding at least one such region is thus  $1 - e^{-\mu(T)} \approx \mu(T)$  when  $\mu(T)$  is small. The approximation becomes asymptotically exact when the number of sib pairs is large and the threshold  $T$  is high. In fact, it is accurate enough for practical purposes provided that  $n$  is at least 50. (It is worth noting that, while lod score analysis of a small number of large families can be quite sensitive to changes in a few key data points, non-parametric statistics based on a large number of small families tend quickly to normal distributions and tend to be robust.)

Suppose, for example, that a study of 100 sib pairs reveals an allele sharing proportion of 61% somewhere in the genome. This result corresponds to a Z-score of 3.1, a lod score of 2.1, and a nominal P value of 0.001. Should one be impressed by this finding? It clearly depends on how often such deviations would arise by chance in a whole genome search.

The mathematical theory of large deviations holds the answer, as was pointed out a few years ago<sup>8,9</sup>. The expected number of chromosomal regions in which a linkage statistic exceeds a threshold  $T$  is given by a simple formula  $\mu(T)$ , explained in Box 1. In fact, the number of such regions approximately follows a Poisson distribution with mean  $\mu(T)$ , and the chance of finding at least one such region is thus  $1 - e^{-\mu(T)} \approx \mu(T)$  when  $\mu(T)$  is small. The approximation becomes asymptotically exact when the number of sib pairs is large and the threshold  $T$  is high. In fact, it is accurate enough for practical purposes provided that  $n$  is at least 50. (It is worth noting that, while lod score analysis of a small number of large families can be quite sensitive to changes in a few key data points, non-parametric statistics based on a large number of small families tend quickly to normal distributions and tend to be robust.)

Suppose, for example, that a study of 100 sib pairs reveals an allele sharing proportion of 61% somewhere in the genome. This result corresponds to a Z-score of 3.1, a lod score of 2.1, and a nominal P value of 0.001. Should one be impressed by this finding? It clearly depends on how often such deviations would arise by chance in a whole genome search.

The mathematical theory of large deviations holds the answer, as was pointed out a few years ago<sup>8,9</sup>. The expected number of chromosomal regions in which a linkage statistic exceeds a threshold  $T$  is given by a simple formula  $\mu(T)$ , explained in Box 1. In fact, the number of such regions approximately follows a Poisson distribution with mean  $\mu(T)$ , and the chance of finding at least one such region is thus  $1 - e^{-\mu(T)} \approx \mu(T)$  when  $\mu(T)$  is small. The approximation becomes asymptotically exact when the number of sib pairs is large and the threshold  $T$  is high. In fact, it is accurate enough for practical purposes provided that  $n$  is at least 50. (It is worth noting that, while lod score analysis of a small number of large families can be quite sensitive to changes in a few key data points, non-parametric statistics based on a large number of small families tend quickly to normal distributions and tend to be robust.)

Suppose, for example, that a study of 100 sib pairs reveals an allele sharing proportion of 61% somewhere in the genome. This result corresponds to a Z-score of 3.1, a lod score of 2.1, and a nominal P value of 0.001. Should one be impressed by this finding? It clearly depends on how often such deviations would arise by chance in a whole genome search.

standard deviations by which  $\pi(x)$  exceeds its null expectation of 0.50; it follows a normal distribution when  $n$  is large. The lod score (or MLS — Maximum Lod Score, the lod score maximized over a set of parameters) is the log-likelihood ratio of the data under the hypothesis that the allele sharing proportion has the observed value  $\pi(x)$  as compared to the hypothesis that there is no excess sharing; the distribution of this statistic is related to a chi-squared distribution when  $n$  is large. The P value reflects the pointwise chance of observing a deviation as high as  $\pi(x)$  under independent assortment.

Suppose, for example, that a study of 100 sib pairs reveals an allele sharing proportion of 61% somewhere in the genome. This result corresponds to a Z-score of 3.1, a lod score of 2.1, and a nominal P value of 0.001. Should one be impressed by this finding? It clearly depends on how often such deviations would arise by chance in a whole genome search.

The mathematical theory of large deviations holds the answer, as was pointed out a few years ago<sup>8,9</sup>. The expected number of chromosomal regions in which a linkage statistic exceeds a threshold  $T$  is given by a simple formula  $\mu(T)$ , explained in Box 1. In fact, the number of such regions approximately follows a Poisson distribution with mean  $\mu(T)$ , and the chance of finding at least one such region is thus  $1 - e^{-\mu(T)} \approx \mu(T)$  when  $\mu(T)$  is small. The approximation becomes asymptotically exact when the number of sib pairs is large and the threshold  $T$  is high. In fact, it is accurate enough for practical purposes provided that  $n$  is at least 50. (It is worth noting that, while lod score analysis of a small number of large families can be quite sensitive to changes in a few key data points, non-parametric statistics based on a large number of small families tend quickly to normal distributions and tend to be robust.)

Fig. 1 shows the results in graphical form. Focusing on P values, we expect regions significant at  $P = 0.05$  to occur about two dozen times by chance (that is, at least once on most chromosomes);  $P = 0.01$  about 7–8 times;  $P = 0.001$  slightly more than once;  $P = 0.0001$  about 0.2 times; and  $P = 0.00002$  about 0.05 times. In other words, there is a 5% chance of randomly finding a region with a P value as extreme as  $2 \times 10^{-5}$ . To keep the chance of encountering a false positive at no more than 5%, one must therefore impose a threshold of  $Z \geq 4.1$ ,  $\text{lod} \geq 3.6$  or  $P = 2 \times 10^{-5}$ . With any less stringent

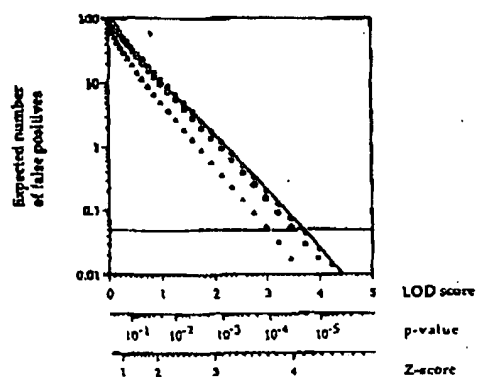


Fig. 1 Number of false positives expected in a whole genome scan for a given threshold of lod score, Z score or pointwise P value. Solid line represents asymptotic expectation for a perfect genetic map, based on the theory described in the Box 1. Symbols represent results for 100 sib pairs obtained from 100,000 simulations using genetic maps with markers spaced every 0.1 cM (circles), every 1 cM (squares), and every 10 cM (triangles). The genome is assumed to consist of 23 chromosomes, with total length 3450 cM. Note the close correspondence between the asymptotic theory and the 0.1 cM simulation. The dotted line indicates the 5% genome-wide significance level.

**Box 2 A simulated genome scan**

To illustrate the standard threshold for a whole-genome scan, we generated simulated data assuming independent assortment of all loci. The simulated data were generated using a computer program that simulates the results of a genome scan. The program generates a set of simulated data for a given number of markers and a given number of chromosomes. The program then calculates the multipoint lod scores for each marker and compares them to the standard threshold of 3.0. The results of the simulation are shown in Box 2.

The standard threshold of 3.0 for a whole-genome scan is based on the assumption of independent assortment of all loci. This assumption is not realistic, as there is a substantial chance (> 5%) of reporting false linkages. To illustrate the point, we describe a simulated whole-genome scan in Box 2.

The standard may seem harsh at first glance, but it accords well with historical practice. The traditional threshold of lod 3 for classical two-point linkage studies of simple mendelian traits corresponds to an asymptotic pointwise significance level of  $P = 10^{-4}$  (ref. 10), for a genome-wide significance level of about 9%. In fact, the lod score threshold need only to be raised to 3.3, corresponding to  $P = 5 \times 10^{-5}$ , to achieve the recommended genome-wide significance level of 5%.

It is worth noting that there is a widespread misconception in human genetics that a lod score of 3 is equivalent to a significance level of only  $P = 10^{-3}$ . This error is rooted in a confusion about the meaning of lod scores and  $P$  values. Lod scores concern the ratio of two probabilities, while  $P$  values refer to a single absolute probability. Specifically, lod = 3 means that the observed data is  $10^3$ -fold more likely to arise under a specified hypothesis of linkage than under the null hypothesis of independent assortment. By contrast,  $P = 10^{-3}$  means that the probability of encountering as large a lod score as observed is  $10^{-3}$  under the null hypothesis. One can convert a lod score to a chi-squared statistic by multiplying by  $2(\log 10) \approx 4.6$  and then use stan-

dard statistical tables, taking into account the one-sided nature of the test, to confirm that a lod score of 2.1 corresponds to  $P = 10^{-3}$ , while a lod score of 3.0 corresponds to the more extreme  $P = 10^{-4}$ .

**Are whole-genome thresholds overly stringent?**

Some geneticists might object to imposing such a stringent standard for declaring linkage. Certain arguments have been advanced in the hopes of gaining special dispensation. It is worth considering them in turn.

• "My study only looked at a few markers (or a few chromosomal regions), so it's not fair to impose a threshold based on a whole genome search." The extreme example of this argument would be a geneticist who finds a weakly positive score with the first marker and seeks to employ the pointwise significance level—asserting that only a single hypothesis has actually been tested. The fallacy is that the investigator would not have abandoned the search if the first marker had been negative, but would have persevered until a positive result was obtained or the entire genome was examined. Having assembled a large patient collection, the geneticist is committed to a whole genome search. It makes no sense to employ a different threshold depending on whether the inevitable false positive fluctuations happen to occur earlier rather than later in the search.

• "My study only involved a genome scan with markers every 10 cM, so it's not fair to impose a threshold based on an infinitely dense genetic map." Again, the analysis does not stop with the sparse map. Initial hints of linkage with a single marker are immediately pursued by using multipoint methods and by peppering the region with a dense collection of markers. In any region that matters, geneticists rapidly extract the complete inheritance information—with the explicit hope of increasing the linkage score.

A hierarchical search—in which one performs a genome scan with a sparse map and then follows up 'interesting' regions with a denser map—is an efficient study design<sup>11, 12</sup>, but the resulting false posi-

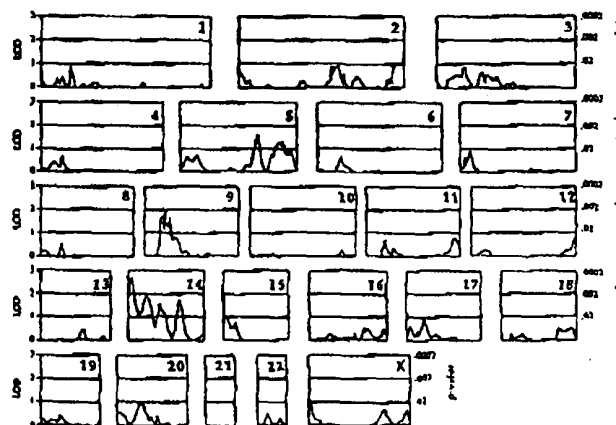


Fig. 2 Simulated genome scan with no trait loci segregating. Chromosomal size is proportional to genetic length, taken from ref. 33. Multipoint lod scores were computed as described<sup>34</sup>.

## progress

tive rate is essentially the same as if a dense map had been used throughout the genome (D. Siegmund, personal communication). This is because the false positives are almost invariably included among the regions chosen for follow-up.

The dense-map threshold turns out not to be that draconian. If one performed only single point analysis with an evenly spaced map, the thresholds for a genome-wide significance level of 5% would not be dramatically different: the lod score thresholds would decrease by ~20% for a 10 cM map; ~15% for a 5 cM map; ~10% for a 2 cM map; and ~7% for a 1 cM map (Fig. 1 shows first and last cases). Moreover, these thresholds would be appropriate only if one did not use multipoint analysis or a denser map to obtain more information. To our thinking, it is better to extract the full inheritance information and find the best *P* value.

In the modern world, it is fair to assume that highly motivated investigators squeeze as much information as possible from the available family material and their results should thus be measured against the corresponding threshold for a dense genome scan. (Some backsliding might be countenanced if strong prior evidence exists to restrict the search to a region; possible cases include a true single-point test of a highly relevant candidate gene, a test of the HLA region for an autoimmune disease, and an X-chromosome scan for a trait with convincing prior evidence of sex linkage.)

Notwithstanding our desire to avoid spurious linkages, we must always remember that regions that fall short of statistical significance may nonetheless be correct. Unfortunately, there is no way to distinguish between small peaks that represent weak true positives and peaks of the same height arising from random fluctuations; assuming that all inheritance information has been extracted. It would be irresponsible to consign such potentially valuable hints to the dustbin of laboratory history. What then is to be done?

## Proposed standards

Back in the days when linkage studies of even the simplest trait required heroic efforts and good fortune, the human genetics community adopted standards to pro-

more both rigor and communication. Lod scores of 3 were required to declare linkage in official chromosome committee reports, but weaker evidence could still be shared in more informal vehicles such as the *McKusick Newsletter*, a predecessor to the modern *Mendelian Inheritance in Man*<sup>13</sup>.

Clear thinking about complex traits would be served by reviving such an approach. Specifically, we propose the following classification based on the number of times that one would expect to see a result at random in a dense, complete genome scan:

- **Suggestive linkage**—statistical evidence that would be expected to occur one time at random in a genome scan.

- **Significant linkage**—statistical evidence expected to occur 0.05 times in a genome scan (that is, with probability 5%).

- **Highly significant linkage**—statistical evidence expected to occur 0.001 times in a genome scan.

- **Confirmed linkage**—significant linkage from one or a combination of initial studies that has subsequently been confirmed in a further sample, preferably by an independent group of investigators. For confirmation, a nominal *P* value of 0.01 should be required (see below).

In the case of sib pair studies, the first three categories would correspond to pointwise significance levels of  $7 \times 10^{-4}$ ,  $2 \times 10^{-5}$ , and  $3 \times 10^{-7}$  and lod scores of 2.2, 3.6, and 5.4. The corresponding *P* values for other study designs differ somewhat (Box 1, Table 1).

Suggestive linkage results will often be wrong, but they are worth reporting—if accompanied by an appropriate warning label about their tenuous nature. Investigators concerned about coming up empty-handed in a genome scan can take comfort from the fact that they can expect, by definition, to find about one suggestive linkage for every trait studied. On the other hand, journal editors must weigh how much attention to accord such results. At the least, specialty journals should actively support the reporting of suggestive linkages in some format. Indeed, it is worth reporting all regions with a nominal *P* value of  $P = 0.05$  encountered in a complete genome scan, but without any claims of linkage.

Because suggestive linkages are so speculative, they should not be assigned gene names lest medical genetics be overrun with illusory loci. Geneticists should enter into a non-proliferation pact, under which genes symbols are reserved for significant linkages. Indeed, traditional usage has been to assign gene names only to confirmed linkages. The appropriate nomenclature committees should take up this issue and develop specific guidelines.

It is worth pointing out that even significant linkages will turn out to be false positives 5% of the time, that is, once in 20 genome scans. Because of the bias that only positive results tend to get reported, the observed false positive rate will be

Table 1 Thresholds for mapping lod underlying complex traits

Mapping method	crossover rate <i>p</i>	suggestive linkage <i>P</i> value (lod)	significant linkage <i>P</i> value (lod)
lod score analysis in human	1	$1.7 \times 10^{-3}$ (1.9)	$4.9 \times 10^{-4}$ (3.3)
Allele-sharing methods in human			
sibs and half-sibs	2	$7.4 \times 10^{-4}$ (2.2)	$2.2 \times 10^{-6}$ (3.6)
grandparent-grandchild	1	$1.7 \times 10^{-3}$ (1.9)	$4.9 \times 10^{-5}$ (3.3)
uncle-nephew	5/2	$5.6 \times 10^{-4}$ (2.3)	$1.8 \times 10^{-6}$ (3.7)
first cousin	8/3	$5.2 \times 10^{-4}$ (2.3)	$1.6 \times 10^{-6}$ (3.7)
first cousin, once removed	20/7	$4.8 \times 10^{-4}$ (2.4)	$1.5 \times 10^{-6}$ (3.8)
second cousin	16/5	$4.2 \times 10^{-4}$ (2.4)	$1.3 \times 10^{-6}$ (3.8)
QTL mapping in mouse or rat			
Backcross (1 d.f.)	1	$3.4 \times 10^{-3}$ (1.9)	$1.0 \times 10^{-4}$ (3.3)
intercross (1 d.f., additive)	1	$3.4 \times 10^{-3}$ (1.9)	$1.0 \times 10^{-4}$ (3.3)
intercross (1 d.f., recessive)	4/3	$2.4 \times 10^{-3}$ (2.0)	$7.2 \times 10^{-4}$ (3.4)
intercross (1 d.f., dominant)	4/3	$2.4 \times 10^{-3}$ (2.0)	$7.2 \times 10^{-4}$ (3.4)
intercross (2 d.f.)	1.5	$1.8 \times 10^{-3}$ (2.0)	$5.2 \times 10^{-5}$ (4.3)

Sib pair analysis involves no dominance component, and thus each sib pair is equivalent to two half-sib pairs. Lod score thresholds for the possible triangle method for sib pairs<sup>37</sup> may be computed by similar methods (D. Siegmund, personal communication); these thresholds are 2.6 for suggestive and 4.0 for significant linkage. Genome size is assumed to be 3300 cM for the human and 1600 cM for the mouse and the rat. A typographical error appeared in the table of ref. 38, which listed the significant *P* value for half-sib and sib pairs as  $3 \times 10^{-6}$ . The correct *P* value is  $2.2 \times 10^{-6}$ , as shown above.

higher in the published literature. While individual investigators cannot do anything about this problem, it offers an additional rationale for conservative standards.

Thomson<sup>14</sup> recently proposed criteria for putative linkage that turn out to be essentially equivalent to our standard for suggestive linkage. Unfortunately, these criteria have been widely misinterpreted as implying genome-wide significance, despite Thomson's clear statement to the contrary. In fact, Thomson (pers. comm.) endorses the standards proposed above.

#### Replications and extensions

Linkage results must be replicated to be credible. We suggest that the term "replication study" should be reserved for situations in which significant linkage has already been obtained in an initial study (or combination of studies). Weaker findings do not merit the same standing as prior hypotheses. For example, there will be many regions with a nominal  $P$  value of 0.05 and some will appear to be 'replicated' in a second study just by chance (Box 2). We prefer the term "extension study" for the process of testing of additional families in the hope of first reaching the genome-wide significance level. Once significant linkage is found, it is appropriate to speak of 'replicating' the result.

Because replication involves testing an established prior hypothesis, the multiple testing problem associated with genome-wide search does not apply. Nonetheless, some caution is still required. The initial localization for a linkage is typically spread over a broad region of about 20 cM. Because one is searching over an interval, there is a multiple testing problem writ small: the chance of finding a  $P$  value of 0.05 somewhere within a 20 cM interval is greater than 5%. It turns out that a pointwise  $P$  value of  $\approx 0.01$  is needed for an interval-wide significance level of 5%. Accordingly,  $P = 0.01$  should be required to declare confirmation at the 5% level. Note that this correction is equivalent to a multiple testing (or Bonferroni) correction for 5 markers.

Failure to replicate does not necessarily disprove a hypothesis. Linkages will often involve weak effects, which may turn out to be weaker in a second study. Indeed, there is a subtle but systematic reason for this: positive linkage results are somewhat biased because they include those weak effects that random fluctuations helped push above threshold, but exclude slightly stronger effects that random fluctuations happened to push below threshold. Initial positive reports will thus tend to overestimate effects, while subsequent studies will regress to the true value (see also ref. 15). Replication studies should always state their power to detect the proposed effect with the given sample size. Negative results are meaningful only if the power is high. Regrettably, many reports neglect this issue entirely.

When several replication studies are carried out, the results may conflict — with some studies replicating the original findings and others failing to do so. This may reflect population heterogeneity, diagnostic differences, or simply statistical fluctuation. Careful meta-analysis of all studies may be useful to assess whether the overall evidence for linkage is convincing.

Suggestive linkages should be pursued in extension studies, in which old and new datasets are combined to see whether a significance evidence of linkage can

be found. To combine results among studies, it is always best to pool the raw data and re-analyze the entire dataset. Lod scores can be added across studies, but only when they are computed by the same method, with the same set of markers, and at the same map position. Other meta-analysis techniques exist<sup>16</sup>.

Statistical aficionados may recognize that extension studies involve a subtle multiple-testing problem of their own, because a significant result in any of the individual datasets or the combined dataset is often taken as evidence of linkage. A modest multiple-testing correction to the genome-wide significance level should therefore be used in extension studies; the appropriate correction depends on study design. Of course, any combined analysis should include all studies — both positive and negative — to avoid biasing the results. If the combined analysis yields significant linkage, it is then appropriate to undertake a replication study.

#### Pursuing hunches

Formal procedures are useful for standardizing the general acceptance of linkage claims. Still, gene hunters should not be inhibited from pursuing all hints and hunches, including: following up all regions with pointwise  $P$  values of 0.05 (even though many will prove to be illusory); being encouraged if they find substantially more suggestive linkages than the one expected by chance (even though real loci cannot be distinguished from false positives); and using epidemiological arguments to infer the existence of loci with small effects (even though such inferences are highly model-dependent). It will, however, be worth having rigorous evidence in hand before undertaking positional cloning to avoid the unpleasant prospect of chasing a phantom locus.

Hints of linkage are usually followed by testing for linkage in larger datasets. Some true susceptibility loci, however, may never show significant linkage because they confer a very small increased risk and have common alleles. The proof that such loci are involved in disease aetiology must come from other data. Linkage disequilibrium can offer a powerful complement to traditional linkage studies. For loci having small effects but relatively few alleles in a population, tests of linkage disequilibrium can be much more sensitive than tests of linkage. A good example is *IDDM2*, the insulin gene, for which strong evidence of linkage disequilibrium is obtained in many datasets that fail to show linkage<sup>17,18</sup>.

Linkage disequilibrium can be used in an exploratory fashion to pursue suggestive (or weaker) linkages (for example, ref. 19). Appropriate correction for multiple testing is essential in such applications — because multiple regions and many different haplotypes are tested for disease association. This topic has not received adequate attention and is an important area for future statistical research. Finally, we note that linkage disequilibrium studies should use family-based controls whenever possible to avoid false positive findings due to population stratification<sup>20-22</sup>.

#### Other models and difficulties

The basic principles above apply to any analysis of

## progress

complex traits — whether by linkage analysis, allele-sharing methods, or quantitative trait mapping in experimental crosses (Box 1). The pairwise  $P$  values vary somewhat according to the method, but they are typically in the range of  $10^{-2}$ – $10^{-4}$  for suggestive and  $10^{-4}$ – $10^{-5}$  for significant linkage.

Nettle problems remain, however. Investigators often try out multiple diagnostic schemes for defining affection status, as well as multiple models of inheritance for linkage analysis. Similarly, studies of quantitative traits may examine a large number of phenotypes. Datasets are frequently stratified using additional criteria, for example HLA genotype. What statistical price should be exacted for such fishing over multiple models? If the models are statistically independent, the observed  $P$  values should be multiplied by the number of models (which is known as the Bonferroni correction). This prescription is too conservative in the case of closely related models (such as correlated phenotypes), but there is no general guidance for how to proceed other than simulation. Even simulation poses a challenge, in that millions of simulations are needed to accurately estimate  $P$  values in the range of  $10^{-5}$ . Techniques such as importance sampling can make simulations much more feasible<sup>23</sup>, and they should be performed whenever possible.

An additional difficulty is that false positive rates can be much higher than estimated if model parameters (such as gene frequency) are misspecified, if sample size is small, and if other assumptions of statistical independence are violated. A careful consideration of all these factors is beyond the scope of this commentary, but they offer an additional reason for caution in interpreting linkage results.

### Examples of complex trait analyses

**IDDM.** Recent genome scans for insulin dependent diabetes mellitus (IDDM) illustrate the issues well. Davies and colleagues<sup>24</sup> used markers at an average spacing of 10 cM to survey the genome in 96 sib pairs, and then followed up some regions with lod  $\geq 1$  ( $P = 0.05$ ) in two further collections with 102 and 84 sib pairs. Sib pairs were analysed together, and also divided according to HLA sharing. In the initial screen, only HLA met the standard for significant linkage, with lod = 8. Two further regions, on chromosomes 8q and X, showed suggestive linkage. A total of 20 regions had lod  $\geq 1$ , which is not significantly greater than would be expected by chance.

Two regions that fell somewhat short of the criterion for suggestive linkage were chosen for followup. A region on chromosome 11q (named *IDDM1* near *FGF*) had a  $P$  value of 0.01 in the combined dataset, but showed suggestive linkage in sib pairs sharing 1 or 0 alleles at HLA. In fact, an independent study found a nearly significant linkage in this region, but only in the subset of sibs in which both carried HLA-DR3<sup>25</sup>. Although the two studies were not jointly analyzed, it is likely the combined data would reveal significant linkage — indicating that this locus is probably real. The second region on chromosome 6q (named *IDDM5*) fell short of suggestive linkage in the combined dataset; it remains unclear whether there is in fact a susceptibility locus in this region. Interestingly,

the same group subsequently reported that a region on chromosome 2 that fell far short of suggestive linkage ( $P = 0.01$ ) showed evidence of linkage disequilibrium in some populations<sup>10</sup>. If widely confirmed (see, for example, ref. 26), this would underscore the value of linkage disequilibrium studies for identification of weak susceptibility loci.

A major contribution of these studies is that they demonstrate that there are no other loci with major effects comparable to HLA. The authors recognized this fact, but provided a valuable spur to further investigation by identifying the most promising regions for further study.

The IDDM story remains a work in progress<sup>27,28</sup>. It will probably require joint analyses of multiple datasets to sort out which of the hints of linkage are real. In general, it would be valuable if data from published genome scans were routinely deposited in an accessible form to facilitate such work.

**Schizophrenia.** Evidence for a susceptibility locus on chromosome 6p in a large collection of pedigrees from Ireland was reported by Wang *et al.*<sup>29</sup> in a recent issue and Straub *et al.*<sup>30</sup> in the current issue of this journal. The lod scores in these papers came extremely close to the standard for significant linkage (corresponding to genome-wide significance levels in the range of 0.05–0.10). The authors carefully stress the need for replication.

Happily, two independent datasets reported in this issue by Moises *et al.*<sup>31</sup> and Schwab *et al.*<sup>32</sup> appear to provide such evidence — each showing suggestive (and nearly significant) linkages in roughly the same region. Although a joint analysis has not yet been undertaken, it is clear that chromosome 6p meets the standard for significant — and, probably, confirmed — linkage.

### Conclusion

The study of complex traits promises to be among the most important and challenging areas of mammalian genetics. As with any endeavor, the field will be shaped by the standards adopted by its practitioners. The traditional threshold of lod 3 has provided a rigorous standard that must be met to declare linkage for a simple Mendelian trait; it corresponds to a genome-wide false positive rate in the neighborhood of 5%. The proposed standard for significant linkage simply extends this same logic to the situation of complex traits; the required  $P$  values accord well with the practice in human genetics over the past three decades. At the same time, the category of suggestive linkage should facilitate reporting of tantalizing but unproven findings. By adopting clear rules for communication, human geneticists will be well prepared for the avalanche of information about to descend.

### Acknowledgments

We thank M. Boehnke, A. Chakravarti, R. Elston, W. Ewens, S. Ghosh, J. MacCluer, M. Mahtani, M. McCarthy, J. Nolan, J. Ott, N. Schork, D. Siegmund, R. Spelman, J. Terwilliger, G. Thomson, J. Todd, and D. Weeks for helpful comments on the manuscript. This work was supported in part by grants from the National Center for Human Genome Research (to E.S.L. and L.K.).



1. Botstein, D., White, D.L., Skolnick, M. & Davis, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. hum. Genet.* 32, 314-331 (1980).
2. Baron, M. et al. Genetic linkage between X-chromosome markers and bipolar affective illness. *Nature* 326, 288-292 (1987).
3. Egeland, J.A. et al. Bipolar affective disorder linked to DNA markers on chromosome 11. *Nature* 326, 783-787 (1987).
4. Kelsoe, J.R. et al. Re-evaluation of the linkage relationship between chromosome 11p loci and the gene for bipolar affective disorder in the Old Order Amish. *Nature* 342, 236-243 (1989).
5. Sherrington, P. et al. Localization of a susceptibility locus for schizophrenia on chromosome 5. *Nature* 338, 164-167 (1988).
6. Kennedy, J.L. et al. Evidence against linkage of schizophrenia to markers on chromosome 5 in a northern Swedish pedigree. *Nature* 336, 167-170 (1988).
7. Baron, M. et al. Diminished support for linkage between manic depressive illness and X-chromosome markers in three Israeli Pedigrees. *Nature Genet.* 3, 49-55 (1993).
8. Lander, E.S. & Botstein, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121, 185-199 (1989).
9. Feingold, E., Brown, P.O. & Sigmund, D. Gaussian models for genetic linkage analysis using complete high-resolution maps of identity by descent. *Am. J. hum. Genet.* 53, 234-251 (1993).
10. Chotai, J. On the lod score method in linkage analysis. *Ann. hum. Genet.* 48, 359-376 (1984).
11. Elston, R.C. Designs for the global search of the human genome by linkage analysis. In *Proceedings of the 15th International Biometrics conference* 39-51 (Hamilton, New Zealand, 1992).
12. Brown, D.L., Golin, M.B. & Weeks, D.E. Efficient strategies for genomic searching using the affected-pedigree-member method of linkage analysis. *Am. J. hum. Genet.* 54, 544-552 (1994).
13. McLaughlin, V.A. & Edwards, J.H. Unassigned syntenic groups and theoretical considerations. Second international workshop on human gene mapping. *Cytogenet. Cell Genet.* 14, 186-198 (1975).
14. Thomson, G. Identifying complex disease genes: progress and paradigms. *Nature Genet.* 8, 108-110 (1994).
15. Suarez, B.K., Hampo, C.L. & Van Eerdewegh, P. Problems of replicating linkage claims in psychiatry. In *Genetic approaches to mental disorders* (eds. Gershon, E.S. & Cloninger, C.R.) 23-46 (American Psychiatric Association, Washington, DC, 1994).
16. Cox, D.R. & Hinkley, D.V. *Theoretical Statistics* (Chapman & Hall, 1974).
17. Spielman, R.S., McGinnis, R.E. & Ewens, W.J. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. hum. Genet.* 52, 605-616 (1993).
18. Bonnes, S.T. et al. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nature Genet.* 9, 284-292 (1995).
19. Cooperman, J.B. et al. Linkage disequilibrium mapping of a type 1 diabetes susceptibility gene (IDDM7) to chromosome 2q31-q33. *Nature Genet.* 8, 80-85 (1995).
20. Ewens, W.J. & Spielman, R.S. The transmission disequilibrium test: history, subdivision, and admixture. *Am. J. hum. Genet.* 57, 455-464 (1995).
21. Gough, S.C.L. et al. Mutation of the glucagon receptor gene and diabetes mellitus in the UK: association or founder effect? *Hum. mol. Genet.* 4, 1609-1612 (1995).
22. Thomson, G. Mapping disease genes: family-based association studies. *Am. J. hum. Genet.* 57, 487-496 (1995).
23. Terwilliger, J.D. & Ott, J. A multi-sample bootstrap approach to the estimation of maximized-over-models 'lod' score distributions. *Cytogenet. Cell Genet.* 58, 142-144 (1992).
24. Davies, J.L. et al. A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 371, 130-136 (1994).
25. Hashimoto, L. et al. Genetic mapping of a susceptibility locus for insulin-dependent diabetes mellitus on chromosome 11q. *Nature* 371, 161-164 (1994).
26. Luo, D., MacLaren, N.K., Huang, H., Muir, A. & She, J. Intrifamilial and case-control association analyses of D2S152 in insulin-dependent diabetes mellitus. *Autoimmunity* (1995, in press).
27. Todd, J.A. Genetic analysis of type 1 diabetes using whole genome approaches. *Proc. natn. Acad. Sci. U.S.A.* 92, 8560-8565 (1995).
28. Luo, D. et al. Affected-sib-pair mapping of a novel susceptibility gene to insulin-dependent diabetes mellitus (IDDM5) on chromosome 6q25-q27. *Am. J. hum. Genet.* 57, 911-919 (1995).
29. Wang, S. et al. Evidence for a susceptibility locus for schizophrenia on chromosome 6p11-p22. *Nature Genet.* 10, 41-46 (1995).
30. Strub, R.E. et al. A potential vulnerability locus for schizophrenia on chromosome 6p24-22: evidence for genetic heterogeneity. *Nature Genet.* 11, 287-293 (1995).
31. Moises, M.W. et al. An international two-stage genome-wide search for schizophrenia susceptibility genes. *Nature Genet.* 11, 321-324 (1995).
32. Schwab, S.G. et al. Evaluation of a susceptibility gene for schizophrenia on chromosome 6p by multipoint affected sib-pair linkage analysis. *Nature Genet.* 11, 325-327 (1995).
33. Gyapay, G. et al. The 1993-94 Genethon human genetic linkage map. *Nature Genet.* 7, 246-339 (1994).
34. Kruglyak, L. & Lander, E.S. Complete multipoint sib pair analysis of qualitative and quantitative traits. *Am. J. hum. Genet.* 57, 439-454 (1995).
35. Kruglyak, L. & Lander, E.S. A nonparametric approach for mapping quantitative trait loci. *Genetics* 139, 1421-1428 (1995).
36. Feingold, E. Markov processes for modeling and analyzing a new genetic mapping method. *J. appl. Prob.* 30, 766-779 (1993).
37. Holmans, P. Asymptotic properties of affected-sib-pair linkage analysis. *Am. J. hum. Genet.* 52, 362-374 (1993).
38. Lander, E.S. & Schork, N.J. Genetic dissection of complex traits. *Science* 263, 2067-2048 (1994).

ED GINN S

## ARTICLES

## Genetic Dissection of Complex Traits

Eric S. Lander\* and Nicholas J. Schork

Medical genetics was revolutionized during the 1980s by the application of genetic mapping to locate the genes responsible for simple Mendelian diseases. Most diseases and traits, however, do not follow simple inheritance patterns. Geneticists have thus begun taking up the even greater challenge of the genetic dissection of complex traits. Four major approaches have been developed: linkage analysis, allele-sharing methods, association studies, and polygenic analysis of experimental crosses. This article synthesizes the current state of the genetic dissection of complex traits—describing the methods, limitations, and recent applications to biological problems.

Human genetics has sparked a revolution in medical science on the basis of the seemingly improbable notion that one can systematically discover the genes causing inherited diseases without any prior biological clue as to how they function. The method of genetic mapping, by which one compares the inheritance pattern of a trait with the inheritance patterns of chromosomal regions, allows one to find where a gene is without knowing what it is. The approach is completely generic, being equally applicable to spongiform brain degeneration as to inflammatory bowel disease.

To geneticists, this revolution is really nothing new. Genetic mapping of trait-causing genes to chromosomal locations dates back to the work of Sturtevant in 1913 (1). It has been a mainstay of experimental geneticists who study fruit flies, nematode worms, yeast, and maize and who developed genetic maps containing hundreds of genetic markers that made it possible to follow the inheritance of any chromosomal region in a controlled cross. With the advent of recombinant DNA, genetic mapping was carried to its logical conclusion with the development of positional cloning—the isolation of a gene solely on the basis of its chromosomal location, without regard to its biochemical function. Positional cloning was invented by Bender and colleagues, who used it to isolate the bithorax complex in *Drosophila* (2), and it rapidly became a routine technique in flies and worms.

Despite its central role in experimental organisms, genetic mapping hardly figured in the study of humans throughout most of the century. There were two reasons: the lack of an abundant supply of genetic markers with which to study inheritance, and the inability

to arrange human crosses to suit experimental purposes. The key breakthrough was the recognition that naturally occurring DNA sequence variation provided a virtually unlimited supply of genetic markers—an idea first conceived of by Botstein and colleagues for yeast crosses (3) and subsequently for human families (4). With highly polymorphic genetic markers, one could trace inheritance in existing human pedigrees as if one had set up the crosses in the laboratory. These ideas soon led to an explosion of interest in the genetic mapping of rare human diseases having simple Mendelian inheritance. More than 400 such diseases have been genetically mapped in this manner, and nearly 40 have been positionally cloned (5).

Human geneticists are now beginning to explore a new genetic frontier, driven by an inconvenient reality: Most traits of medical relevance do not follow simple Mendelian monogenic inheritance. Such "complex" traits include susceptibilities to heart disease, hypertension, diabetes, cancer, and infection. The genetic dissection of complex traits is attracting many investigators with the promise of shedding light on old problems and is spawning a variety of analytical methods. The emerging issues turn out to be relevant not just to medical genetics, but to fundamental studies of mammalian development and applied work in agricultural improvement. The field is still at an early stage, but it is ready to explode much as it has done in recent years with the analysis of simple traits. The purpose of this article is to synthesize the key challenges and methods, to highlight some enlightening examples, and to identify further needs.

## Complex Traits

The term "complex trait" refers to any phenotype that does not exhibit classic Mendelian recessive or dominant inheritance attributable to a single gene locus. In general, complexities arise when the simple correspondence between genotype and phenotype

breaks down, either because the same genotype can result in different phenotypes (due to the effects of chance, environment, or interactions with other genes) or different genotypes can result in the same phenotype.

To some extent, the category of complex traits is all-inclusive. Even the simplest genetic disease is complex, when looked at closely. Sickle cell anemia is a classic example of a simple Mendelian recessive trait. Yet, individuals carrying identical alleles at the  $\beta$ -globin locus can show markedly different clinical courses, ranging from early childhood mortality to a virtually unrecognized condition at age 50 (6). The trait of severe sickle cell anemia is thus complex, being influenced by multiple genetic factors including a mapped X-linked locus and an inferred autosomal locus that can increase fetal hemoglobin amounts and thereby partially ameliorate the disease (7).

It is often impossible to find a genetic marker that shows perfect cosegregation with a complex trait. The reasons for this can be ascribed to a few basic problems.

**Incomplete penetrance and phenocopy.** Some individuals who inherit a predisposing allele may not manifest the disease (incomplete penetrance), whereas others who inherit no predisposing allele may nonetheless get the disease as a result of environmental or random causes (phenocopy). Thus, the genotype at a given locus may affect the probability of disease, but not fully determine the outcome. The penetrance function ( $f(G)$ ), specifying the probability of disease for each genotype  $G$ , may also depend on nongenetic factors such as age, sex, environment, and other genes. For example, the risk of breast cancer by ages 40, 55, and 80 is 37%, 66%, and 85% in a woman carrying a mutation at the *BRCA1* locus as compared with 0.4%, 3%, and 8% in a noncarrier (8). In such cases, genetic mapping is hampered by the fact that a predisposing allele may be present in some unaffected individuals or absent in some affected individuals.

**Genetic (or locus) heterogeneity.** Mutations in any one of several genes may result in identical phenotypes, such as when the genes are required for a common biochemical pathway or cellular structure. This poses no problem in experimental organisms, because geneticists can arrange to work with pure-breeding strains and perform crosses to assign mutations to complementation classes. In contrast, medical geneticists typically have no way to know whether two patients suffer from the same disease for

E. S. Lander is with the Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA, and the Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. N. J. Schork is with the Department of Genetics and Center for Human Genetics, Case Western Reserve University School of Medicine and University Hospitals of Cleveland, Cleveland, OH 44106, USA.

\*To whom correspondence should be addressed.

different genetic reasons, at least until the genes are mapped. Examples of genetic heterogeneity in humans include polycystic kidney disease (9), early-onset Alzheimer's disease (10), maturity-onset diabetes of the young (11), hereditary nonpolyposis colon cancer (12), ataxia telangiectasia (13), and xeroderma pigmentosum (14). Retinitis pigmentosa, involving retinal degeneration, apparently can result from mutations in any of at least 14 different loci (15), and Zellweger syndrome, involving the failure of peroxisome biosynthesis, from mutations in any of 13 loci (16). Genetic heterogeneity hampers genetic mapping, because a chromosomal region may cosegregate with a disease in some families but not in others. Genetic heterogeneity should be distinguished from allelic heterogeneity, in which there are multiple disease-causing mutations at a single gene. Allelic heterogeneity tends not to interfere with gene mapping.

**Polygenic inheritance.** Some traits may require the simultaneous presence of mutations in multiple genes. Polygenic traits may be classified as discrete traits, measured by a specific outcome (for example, development of type 1 diabetes or death from myocardial infarction), or quantitative traits, measured by a continuous variable (for example, diastolic blood pressure, fasting glucose concentrations, or immunoglobulin E (IgE) titers) whose level may be set by the combined action of individual quantitative trait loci. Discrete traits may represent a threshold effect, produced whenever an underlying quantitative variable, influenced by multiple genes, exceeds a critical threshold, or a pure synthetic effect, requiring the simultaneous and joint action of each of several mutations.

Polygenic inheritance is easily demonstrated in animal crosses, in the transmission pattern of quantitative traits such as blood pressure (17), and in the pervasive "genetic background" effects that represent the action of modifier genes. For example, a mutation in the mouse *Apc* gene causes numerous intestinal neoplasias and early death in B6 mice but has barely noticeable effects when bred into an AKR strain (18). More generally, the phenotype of "knock-out mice" may vary dramatically on different strain backgrounds, pointing to previously unknown interacting genes.

Polygenic inheritance is harder to demonstrate directly in humans, but it is surely no less common. One form of retinitis pigmentosa was shown to be due to strict digenic inheritance, requiring the presence of heterozygous mutations at the peripherin/RDS and *ROM1* genes (19) (whose encoded proteins are thought to interact in the photoreceptor outer segment disc membranes). Some forms of Hirschsprung disease appear to require the simultaneous

presence of mutations on chromosomes 13, 21, and possibly elsewhere (20). Polygenic inheritance complicates genetic mapping, because no single locus is strictly required to produce a discrete trait or a high value of a quantitative trait [except in the case of a pure synthetic interaction causing a discrete trait (21, 22)].

**High frequency of disease-causing alleles.** Even a simple trait can be hard to map if disease-causing alleles *D* occur at high frequency in the population. The expected Mendelian inheritance pattern of disease will be confounded by the problem that multiple independent copies of *D* may be segregating in the pedigree [often referred to as bilineality (23)] and that some individuals may be homozygous for *D* [in which case one will not observe linkage between *D* and a specific allele at a nearby genetic marker, because either of the two homologous chromosomes could be passed to an affected offspring (24)]. Late-onset Alzheimer's disease provides an excellent example. Initial linkage studies found weak evidence of linkage to chromosome 19q, but they were dismissed by many observers because the lod score (logarithm of the likelihood ratio for linkage) remained relatively low, and it was difficult to pinpoint the linkage with any precision (25). The confusion was finally resolved with the discovery that the apolipoprotein E type 4 allele appears to be the major causative factor on chromosome 19. The high frequency of the allele (~16% in most populations) had interfered with the traditional linkage analysis (26). High frequency of disease-causing alleles becomes an even greater problem if genetic heterogeneity is also present.

**Other transmission mechanisms.** Finally, mammalian genetics has revealed additional modes of genetic inheritance. These include mitochondrial inheritance (in which mitochondria pass solely through the maternal germ line, and each meiotic transmission may involve selection from a potentially mixed population of mutant and normal organelles); imprinting (due to differential activity of the paternal and maternal copies of a gene); and phenomena due to the expansion of trinucleotide repeats such as so-called "anticipation." These modes of transmission pose little difficulty when they obey strict rules (as for imprinting), but they can complicate analysis when they lead to highly variable transmission rates [as for some mitochondrial diseases or trinucleotide repeat diseases (27)] and may require specialized methods (28).

## Genetic Epidemiology

Before undertaking DNA-based studies aimed at genetic dissection, one would ideally like to infer as much as possible about the genetic basis of a trait on the basis of

the pattern of disease incidence in families and populations. Such genetic epidemiology constitutes a major field in its own right for which excellent reviews exist (29). We focus on a few key concepts.

**Twin studies.** Whereas experimental geneticists can propagate inbred lines with isogenic genetic constitution, the only opportunity to examine the expression of a human trait in a fixed genetic background comes from the study of monozygotic (MZ) twins (30). The absolute risk to an MZ twin of an affected individual provides a direct estimate of penetrance for a given environment.

**Relative risk.** The most important epidemiological parameter is the relative risk,  $\lambda_R$ , defined as the recurrence risk for a relative of an affected person divided by the risk for the general population. The subscript *R* denotes the type of relation; for example,  $\lambda_O$  and  $\lambda_S$  are the risks to offspring and sibs, respectively. The magnitude of  $\lambda_R$  is related to the degree of concordant inheritance for genetic determinants in affected relative pairs and thus is related to the ease or difficulty of genetic mapping, as shown by Risch (31–33). Genetic mapping is much easier for traits with high  $\lambda$  (for example,  $\lambda_S > 10$ ) than for those with low  $\lambda$  (for example,  $\lambda_S < 2$ ). As an illustration of the range,  $\lambda_S \approx 500$  for cystic fibrosis; 15 for type 1 diabetes [of which a factor of 3 to 4 is attributable to concordance at the human leukocyte antigen (HLA) complex]; 8.6 for schizophrenia; and 3.5 for type II diabetes. For a quantitative phenotype, a similar measure is the heritability of the trait (34).

**Segregation analysis.** Segregation analysis involves fitting a general model to the inheritance pattern of a trait in pedigrees. Using a model involving the presence of a simple Mendelian factor in a background of multifactorial inheritance, one tries to estimate key parameters such as the allele frequency, penetrance, and proportion of cases explained by the Mendelian factor. An important example is the work of Newman et al. and other researchers (35, 36) who showed that the degree of familial clustering for breast cancer observed in 1579 nuclear families was consistent with a dominantly acting rare allele (frequency = 0.06%), accounting for 4% of affected women (but 20% of affected mother-daughter pairs), in a larger background of multifactorial causation. Segregation analysis can be extremely sensitive to biases in the ascertainment of families [for example, if preferential inclusion of affected individuals may cause the penetrance to be greatly overstated (37)], and it may have little ability to distinguish among the many possible modes of inheritance for complex traits (38). Moreover, it can

## ARTICLES

be especially difficult to estimate the number of distinct genes influencing a trait, except in very favorable situations (39), and to identify penetrance parameters associated with multiple loci (40).

## Defining Diseases

Given the many problems that can hamper genetic dissection of complex traits, geneticists try to stack the deck in their favor. By narrowing the definition of a disease or restricting the patient population, it is often possible to work with a trait that is more nearly Mendelian in its inheritance pattern and more likely to be homogeneous. The extent to which redefinition simplifies the task of genetic mapping can be measured by the resulting increase in the relative risk  $\lambda_R$ . Although there is no guaranteed method to increase  $\lambda_R$ , four criteria are often useful.

**Clinical phenotype.** For example, when colon cancer is restricted to cases with extreme polyposis, the trait becomes a simple autosomal dominant one—which allowed positional cloning of the APC gene on chromosome 5 (41). Other forms of colon cancer can be distinguished by the phenotype of replication errors in tumors (42). In studying hypertension, one can increase  $\lambda$  by focusing on cases with combined hypertension and hyperlipidemia (43).

**Age at onset.** Breast cancer and Alzheimer's disease are rendered genetically more homogeneous by focusing on early-onset cases [although the latter can be caused by at least three independent loci (44)]. Similarly, the relative risk for death from heart attack is much greater for early-onset cases ( $\lambda_2 = 7$  in men and  $\lambda_2 = 15$  in women under age 65) as compared with late-onset cases ( $\lambda_2 < 2$ ) (45).

**Family history.** For example, the sister of a woman with breast cancer has a much greater risk if her mother is also affected (35, 36). Hereditary nonpolyposis colon cancer (12) was genetically mapped by defining the trait to require the presence of at least two other affected relatives.

**Severity.** For continuous traits, it often pays to consider as affected only those individuals at the extreme ends of the trait distribution. For example, one might select families for a hypertension study on the basis of the presence of at least one member with blood pressure exceeding 140/90. Such selection can greatly increase the ability to map genes, both in human families (46) and experimental crosses (47).

Another way to improve the prospects for genetic dissection is to focus on specific ethnic groups. Population genetic theory and data suggest that there will be greater genetic and allelic homogeneity in a more genetically isolated group (such as Sardinians, Basques, Finns, and Japanese) than in

a large, mixed population (such as is in New York City, or Los Angeles). Different ethnic groups may shed light on different aspects of a disease, which might be much harder to discern in an outbred population. For example, it has been suggested that there may be differences in the genetic etiology of type II diabetes between Mexican Americans and Scandinavians, with somewhat higher frequency of early insulin resistance in the former and an early pancreatic beta cell defect in the latter (48). Focusing on a highly restricted population may also offer advantages for eventual positional cloning, because one may be able to exploit linkage disequilibrium for fine-structure genetic mapping (discussed below).

Genetic Dissection:  
The Fourfold Way

The methods available for genetic dissection of complex traits fall neatly into four categories: linkage analysis, allele-sharing methods, association studies in human populations, and genetic analysis of large crosses in model organisms such as the mouse, and rat.

## Linkage Analysis

Linkage analysis involves proposing a model to explain the inheritance pattern of phenotypes and genotypes observed in a pedigree (Fig. 1). It is the method of choice for simple Mendelian traits because the allowable models are few and easily tested. However, applications to complex traits can be more problematic, because it may be hard to find a precise model that adequately explains the inheritance pattern.

Formally, linkage analysis consists of finding a model  $M_1$ , positing a specific location for a trait-causing gene, that is much more likely to have produced the observed data than a null hypothesis  $M_0$ , positing no linkage to a trait-causing gene in the region. The evidence for  $M_1$  versus  $M_0$  is measured by the likelihood ratio,  $LR = \text{Prob}$

$(\text{Data}|M_1)/\text{Prob}(\text{Data}|M_0)$ , or, equivalently, by the lod score,  $Z = \log_{10}(LR)$  (49, 50).

The model  $M_1$  is typically chosen from among a family of models  $M(\Phi)$ , where  $\Phi$  is a parameter vector that might specify such information as the location of the trait-causing locus, the allele frequency at the trait and marker loci, the penetrance function, and the transmission frequencies from parent to child. Many of these parameters may already be known (such as penetrance functions from prior segregation analysis or marker allele frequencies from population surveys). The remaining, unknown parameters are chosen to be the maximum likelihood (ML) estimate, that is, the value  $\hat{\Phi}$  that makes the data most likely to have occurred (51). The null model  $M_0$  corresponds to a specific null hypothesis about the parameters,  $\Phi_0$ .

For example, the model for a simple Mendelian recessive or dominant disease is completely specified except for the recombination frequency  $\theta$  between the disease gene and a marker; the null hypothesis of nonlinkage corresponds to  $\theta = 50\%$  recombination.

The ML model  $M(\hat{\Phi})$  is accepted (compared with  $M_0$ ) if the corresponding maximum lod score  $\hat{Z}$  is large, that is, exceeds a critical threshold  $T$ . Of course, a crucial issue is the appropriate significance threshold. The traditional lod score threshold has been 3.0 (50, 52), although the appropriateness of this choice is discussed in the section on statistical significance.

**Applications.** Linkage analysis is the current workhorse of human genetic mapping, having been applied to hundreds of simple monogenic traits. Linkage analysis has also been successfully applied to genetically heterogeneous traits in some cases. The simplest situation is when unequivocal linkage can be demonstrated in a single large pedigree (with  $Z \gg 3$ ), even though other families may show no linkage. This has been done for such diseases as adult polycystic kidney disease, early-onset Alzheimer's disease, and psoriasis (53). If linkage cannot be established on the basis of any single pedigree, one can ask whether a subset of the pedigrees collectively shows evidence of linkage. Of course, one cannot simply choose those families with positive lod scores and exclude those with negative lod scores, as such an *ex post* selection criterion will always produce a high positive lod score. Instead, one must explicitly allow for genetic heterogeneity within the linkage model (through the inclusion of an admixture parameter  $\alpha$  specifying the proportion of linked families), although care is required because the resulting lod score has irregular statistical properties (54). Alternatively, families can be selected on the basis of a priori considerations. An example of this

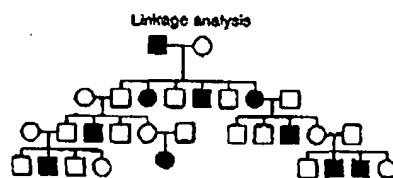


Fig. 1. Linkage analysis involves constructing a transmission model to explain the inheritance of a disease in pedigrees. The model is straightforward for simple Mendelian traits but can become very complicated for complex traits. Linkage analysis has been applied to hundreds of simple Mendelian traits, as well as to such situations as genetic heterogeneity in breast cancer and two-gene interactions in multiple sclerosis.

approach is provided by the genetic mapping of a gene for early-onset breast cancer (BRCA1) to chromosome 17q (55). Families were added to the linkage analysis in order of their average age of onset, resulting in a lod score that rose steadily to a peak of 2 = 6.0 with the inclusion of families with onset before age 47 and then fell with the addition of later-onset pedigrees. Notwithstanding these successes, many failed linkage studies may result from cryptic heterogeneity. It is always wise to try to redefine traits to make them more homogeneous.

Linkage analysis can also be applied when penetrance is unknown. One approach is to estimate the ML value of the penetrance  $p$  within the linkage analysis. A particular concern is to avoid incorrectly overestimating  $p$ , because this can lead to spurious evidence against linkage (caused by individuals who inherit a trait-causing allele but are unaffected). One can guard against this problem by performing an affecteds-only analysis, in which one records unaffected individuals as "phenotype unknown" or, equivalently, sets the penetrance artificially low ( $p \approx 0$ ). This approach was important in studies of both early-onset and late-onset Alzheimer's disease (25, 56). In the latter case, the lod score increased from 2.20 with an age-adjusted penetrance function to 4.38 with an affecteds-only analysis.

Some traits are so murky that it is unclear who should be considered affected. Psychiatric disorders fall into this category, and investigators have explored using various alternative diagnostic schemes within their analysis. For example, schizophrenia might be defined strictly to include only patients meeting the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) criteria or be defined more loosely to include patients with so-called schizoid personality disorders (57). This approach is permissible in theory but requires great care in adjusting the significance level to offset the effect of multiple hypothesis testing.

Linkage analysis can also be extended to situations in which two or more genes play a role in the inheritance of a disease, simply by examining the inheritance pattern of pairs of regions. Such an approach has been dubbed simultaneous search (21, 58, 59). It can be applied to the situation of a genetically heterogeneous trait or to an interaction between two loci. Multiple sclerosis in large Finnish kindreds has been reported to be linked to the inheritance of both HLA on chromosome 6 and the gene for myelin basic protein on chromosome 18, on the basis of such a two-locus analysis (60).

**Limitations.** Linkage analysis is subject to the same limitations as any model-based method. It can be very powerful, provided that one specifies the correct model (61,

62). Use of the wrong model, however, can lead one to miss true linkages and sometimes to accept false linkages (63, 64). In particular, exclusion mapping of regions can only demonstrate absence of a trait-causing locus fitting the particular model tested (50, 52). Finally, testing many models requires the use of a higher significance level, which may decrease the power to detect a gene; this issue is discussed in the section on statistical significance. The more complex the trait, the harder it is in general to use linkage analysis (65).

**Computation.** Calculating the likelihood ratio can be horrendously complicated in some cases and requires computer programs (66, 67). Elston and Stewart invented the first practical algorithm for calculating likelihoods (68, 69), which was implemented by Ott in the first general-purpose linkage program LIPED (70) and is also at the heart of the widely used LINKAGE package (71). However, the algorithm is not a complete panacea. In its original form it does not easily accommodate environmental or polygenic covariation among family members, which form the basis of so-called "mixed models" (67, 72) used widely in genetic epidemiology (73). In addition, it can be extremely slow for analysis with many genetic markers or inbred families. Alternative exact algorithms have been developed for some applications (74), including one that allows multipoint homozygosity mapping (75), but these tend to be limited to smaller pedigrees. Likelihoods can also be estimated by simulation-based methods, such as the Gibb's sampler and Monte Carlo Markov chains (76). There remain many important computational challenges in linkage analysis.

### Allele-Sharing Methods

Allele-sharing methods are not based on constructing a model, but rather on rejecting a model. Specifically, one tries to prove that the inheritance pattern of a chromosomal region is not consistent with random Mendelian segregation by showing that affected relatives inherit identical copies of the region more often than expected by chance (Fig. 2). Because allele-sharing methods are nonparametric (that is, assume no model for the inheritance of the trait), they tend to be more robust than linkage analysis: affected relatives should show excess allele sharing even in the presence of incomplete penetrance, phenocopy, genetic heterogeneity, and high-frequency disease alleles. The tradeoff is that allele-sharing methods are often less powerful than a correctly specified linkage model.

Allele-sharing methods involve studying affected relatives in a pedigree to see how often a particular copy of a chromosomal

region is shared identical-by-descent (IBD), that is, is inherited from a common ancestor within the pedigree. The frequency of IBD sharing at a locus can then be compared with random expectation. Formally, one can define an identity-by-descent affected-pedigree-member (IBD-APM) statistic

$$t(s) = \sum_{i,j} x_{ij}(s)$$

where  $x_{ij}(s)$  is the number of copies shared IBD at position  $s$  along a chromosome, and where the sum is taken over all distinct pairs ( $i, j$ ) of affected relatives in a pedigree. The results from multiple families can be combined in a weighted sum  $T(s)$ . Assuming random segregation,  $T(s)$  tends to a normal distribution with a mean  $\mu$  and variance  $\sigma$  that can be calculated on the basis of the kinship coefficients of the relatives compared (77, 78). Deviation from random segregation is detected when the statistic  $(T - \mu)/\sigma$  exceeds a critical threshold (see the section on statistical significance).

**Sib pairs.** Affected sib pair analysis is the simplest form of this method. For example, two sibs can show IBD sharing for zero, one, or two copies of any locus (with a 25%-50%-25% distribution expected under random segregation). If both parents are available, the data can be partitioned into separate IBD sharing for the maternal and paternal chromosome (zero or one copy, with a 50%-50% distribution expected under random segregation). In either case, excess allele sharing can be measured with a simple  $\chi^2$  test (79-81).

Sib pair studies have played an important role in the study of type 1 diabetes. Excess allele sharing confirmed the important role of HLA, although the inheritance pattern fit neither a simple dominant or recessive model (82, 83). With the availability of a comprehensive human genetic

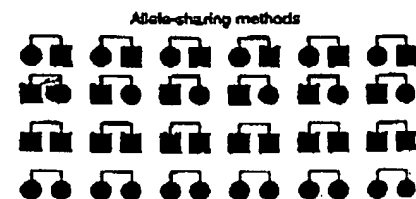


Fig. 2. Allele-sharing methods involve testing whether affected relatives inherit a region identical-by-descent (IBD) more often than expected under random Mendelian segregation. Affected sib pair analysis is a well-known special case, in which the presence of a trait-causing gene is revealed by more than the expected 50% IBD allele sharing. The method is more robust for genetic complications than linkage analysis but can be less powerful than a correctly specified linkage model. Examples include applications to type 1 diabetes, essential hypertension, IgE levels, and bone density in postmenopausal women.

## ARTICLES

(IBD),  
cestor  
of IBD  
pared  
one  
cred-

hared  
:, and  
pairs  
The  
com-  
ran-  
l dis-  
ship  
(77,  
n is  
ex-  
n on

the  
ple,  
me,  
5%  
py,  
in-  
se,  
a

it-  
ss-  
re-  
ce  
or  
l-  
c

linkage map, sib pair analysis has been applied to a whole-genome scan, and excess allele sharing has been found at a locus on chromosome 11q, pointing to a previously unidentified causal factor in type 1 diabetes (34). In a similar search restricted to the X chromosome, brothers concordant for the trait of homosexual orientation showed significant excess allele sharing (33 out of 40 cases) in the region Xq28, suggesting the involvement of a genetic factor influencing at least the particular subtype of homosexuality studied (85). The same approach can be applied to affected uncle-nephew pairs and cousin pairs, for example.

**IBD versus IBS.** One often cannot tell whether two relatives inherited a chromosomal region IBD, but only whether they have the same alleles at genetic markers in the region, that is, are identical by state (IBS). It is usually safe to infer IBD from IBS when a dense collection of highly polymorphic markers has been examined, but the early stages of genetic analysis may involve sparser maps with less informative markers. Two approaches have been developed to cope with this important practical difficulty. The first amounts to inferring IBD sharing on the basis of the marker data (expected IBD-APM methods) (86), whereas the second uses another statistic based explicitly on IBS sharing (IBS-APM method) (78, 87). (The inventors of the latter method dubbed it simply the APM method, but we prefer the more descriptive names used here.) Both approaches are important, although key statistical and computational issues remain open for each.

A number of recent studies have applied IBS-APM methods to complex traits. The angiotensinogen gene has been shown with IBS-APM analysis to be linked to essential hypertension in multiplex families, although the gene explains only a minority of the phenotype (88). Similarly, linkage of late-onset Alzheimer's disease to chromosome 19 could be established by IBS-APM, even though traditional lod score analysis gave more equivocal results (25).

**Quantitative traits.** Allele-sharing methods can also be applied to quantitative traits. An approach proposed by Haseman and Elston (89) is based on the notion that the phenotypic similarity between two relatives should be correlated with the number of alleles shared at a trait-causing locus. Formally, one performs regression analysis of the squared difference  $\Delta^2$  in a trait between two relatives and the number  $x$  of alleles shared IBD at a locus. The approach can be suitably generalized to other relatives (90) and multivariate phenotypes (91). It has been used, for example, to relate serum IgE levels with allele sharing in the region of the gene encoding interleukin-4 and bone density in postmenopausal wom-

en with allele sharing in the region of the vitamin D receptor (92, 93). In addition, there has been a resurgence of interest in the theoretical aspects of mapping genes with IBD and IBS methods (94).

APM methods have been applied to whole-genome searches only in a few cases, including a recent study on manic depression (95). This situation is certain to change in the near future.

### Association Studies

Association studies do not concern familial inheritance patterns at all. Rather, they are case-control studies based on a comparison of unrelated affected and unaffected individuals from a population (Fig. 3). An allele  $A$  at a gene of interest is said to be associated with the trait if it occurs at a significantly higher frequency among affected compared with control individuals. The statistical analysis is simple, involving only a  $2 \times 2$  contingency table. The biggest potential pitfall of association studies is in the choice of a control group (which is in sharp contrast to linkage and allele-sharing methods, which require no control group because they involve testing a specific model of random Mendelian segregation within a family). Although association studies can be performed for any random DNA polymorphism, they are most meaningful when applied to functionally significant variations in genes having a clear biological relation to the trait.

Association studies have played a crucial role in implicating the HLA complex in the etiology of autoimmune diseases. The allele HLA-B27, for example, occurs in 90% of patients with ankylosing spondylitis but only 9% of the general population (96). There are scores of HLA associations involving such diseases as type 1 diabetes, rheumatoid arthritis, multiple sclerosis, celiac disease, and systemic lupus erythematosus (97). More recently, association studies played a key role in implicating the

apolipoprotein E gene in both late-onset Alzheimer's disease and heart disease and the angiotensin converting enzyme (ACE) gene in myocardial infarction (98). In addition, methods for assessing associations between marker loci and quantitative traits have received recent attention (99).

What does a positive association imply about a disease? On its own, very little. Associations can arise for three reasons, one of which is completely artifactual.

1) Positive association can occur if allele  $A$  is actually a cause of the disease. In this case, the same positive association would be expected to occur in all populations (100).

2) Positive association can also occur if allele  $A$  does not cause the trait but is in linkage disequilibrium with the actual cause, that is,  $A$  tends to occur on those chromosomes that also carry a trait-causing mutation. Linkage disequilibrium will arise in a population when two conditions are met: most cases of the trait are due to relatively few distinct ancestral mutations at a trait-causing locus, and the marker allele  $A$  was present on one of these ancestral chromosomes and lies close enough to the trait-causing locus that the correlation has not yet been eroded by recombination during the population's history. Linkage disequilibrium is most likely to occur in a young, isolated population.

True associations due to linkage disequilibrium can yield seemingly contradictory results. Because linkage disequilibrium depends on a population's history, a trait might show positive association with allele  $A_1$  in one isolated population, with allele  $A_2$  in second isolated population, and with no allele in a large, mixed population. Moreover, a trait may show no association with an Eco RI restriction fragment length polymorphism (RFLP) in a gene but strong association with a nearby Bam HI RFLP, because of the particular population genetic features of a population (101).

3) Most disturbingly, positive association can also arise as an artifact of population admixture. In a mixed population, any trait present at a higher frequency in an ethnic group will show positive association with any allele that also happens to be more common in that group. To give a light-hearted example, suppose that a would-be geneticist set out to study the "trait" of ability to eat with chopsticks in the San Francisco population by performing an association study with the HLA complex. The allele HLA-A1 would turn out to be positively associated with ability to use chopsticks—not because immunological determinants play any role in manual dexterity, but simply because the allele HLA-A1 is more common among Asians than Caucasians.

Association studies

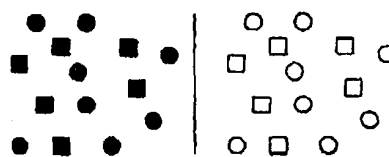


Fig. 3. Association studies test whether a particular allele occurs at higher frequency among affected than unaffected individuals. Association studies thus involve population correlation, rather than cosegregation within a family. Examples include HLA associations in many autoimmune diseases, apolipoprotein E4 in Alzheimer's, and angiotensin converting enzyme (ACE) in heart disease.

This problem has afflicted many association studies performed in inhomogeneous populations ranging from the population of metropolitan Los Angeles to Native American tribes. A subtle example arose because Pima Amerindians are much more susceptible than Caucasians to type II diabetes. Studies in the Pima showed association between type II diabetes and the Gm locus, with the "protective" allele being the one present at higher frequency in Caucasians. Subsequent work, however, revealed that the association was apparently because tribe members have different degrees of Caucasian ancestry. The presence of a "Caucasian" allele at any gene tends to correlate with a higher degree of Caucasian ancestry, which in turn tends to correlate with a lower risk of type II diabetes (102).

To prevent spurious associations arising from admixture, a number of steps should be taken.

1) If possible, association studies should be performed within relatively homogeneous populations. If an association can only be found in large, mixed populations but not in homogeneous groups, one should suspect admixture.

2) Given the difficulty of selecting a control group that is perfectly matched for ethnic ancestry, association studies should use an "internal control" for allele frequencies: a study of affected individuals and their parents. If the parents have genotypes  $A_1/A_2$  and  $A_3/A_4$ , and the affected individual has genotype  $A_1/A_3$ , then the genotype  $A_2/A_4$  (consisting of the two alleles that the affected individual did not inherit) provides an "artificial control" that is well matched for ethnic ancestry. This method is sometimes called the affected family-based control or haplotype relative risk method and can be applied either to the genotypes or to the alleles (103). In our opinion, such internal controls should be routinely used.

Collecting parental DNA is useful for a second, unrelated reason. With knowledge of parental genotypes, one can construct multimarker haplotypes (indicating the alleles found on the same maternally or paternally derived chromosome), which can be much more informative than studying single markers one at a time. This can be especially useful in isolated populations, where only a limited number of distinct trait-causing chromosomes may be present.

3) Once a tentative association has been found, it should be subjected to a transmission disequilibrium test (TDT) (104, 105). The test has the premise that a parent heterozygous for an associated allele  $A_1$  and a nonassociated allele  $A_2$  should more often transmit  $A_1$  than  $A_2$  to an affected child. The TDT was first applied to the puzzling situation of the insulin gene, which showed strong association but no

linkage to type I diabetes; linkage had been obscured because of the substantial proportion of homozygous (and thus non-segregating) parents (104). It should be noted that TDT cannot be directly applied to the sample in which initial association was found (because affected individuals necessarily have an excess of the associated allele) but rather to a new sample from the same population.

The controversy over a reported association between alcoholism and an allele at the dopamine D2 receptor (DRD2) illustrates all the issues in association studies. The initial study compared postmortem samples from 35 alcoholics and 35 controls, with no attempt to control for ethnic ancestry (other than race) (106). For a Taq I RFLP located about 10 kb downstream from DRD2, the A1 allele was found to be present in 69% of alcoholics and 27% of controls. Attempts to replicate this finding, however, have yielded conflicting results, with some authors finding no association whatsoever and others reporting association for severe alcoholism only (107). Revealingly, the frequency of the polymorphism has been shown to vary substantially among populations and among the various "control" groups used. In light of this variation, it is imperative that studies use internal control genotypes, although this has not been done to date. Association studies in relatively homogeneous populations, linkage studies, and transmission tests have all been negative (108). At present, there is no compelling evidence that the reported association is not an artifact of admixture.

Association studies are not well suited to whole-genome searches in large, mixed populations. Because linkage disequilibrium extends over very short distances in an old population (109), one would need tens of thousands of genetic markers to "cover" the genome. Moreover, testing many markers raises a serious problem of multiple hypothesis testing: each association test is nearly independent. Testing  $n$  loci each with  $k$  alleles amounts to performing about  $n(k-1)$  independent tests, and the required significance level should be divided by this factor. A nominal significance level of  $P = 0.0001$  is thus needed simply to achieve an overall false positive rate of 5%, if one tests 100 markers with six alleles each. (Some authors propose to avoid this problem by identifying all results significant at the  $P = 0.05$  level in an initial sample and then attempting to replicate them in a second sample (110). However, the same multiple testing issue still applies to retesting many results at the second stage.) Genomic search for association may be more favorable in young, genetically isolated populations because linkage disequilibrium extends over greater distances, and the num-

ber of disease-causing alleles is likely to be fewer (21, 111).

In summary, linkage-type studies and association studies have many crucial differences. Association studies test whether a disease and an allele show correlated occurrence in a population, whereas linkage studies test whether they show correlated transmission within a pedigree. Association studies focus on population frequencies, whereas linkage studies focus on concordant inheritance. One may be able to detect linkage without association (for example, when there are many independent trait-causing chromosomes in a population, so that association with any particular allele is weak) or association without linkage (for example, when an allele explains only a minor proportion of the variance for a trait, so that the allele may occur more often in affected individuals but does a poor job of predicting disease status within a pedigree). Linkage and association are often used interchangeably in popular articles about genetics, but this practice should always be avoided.

### Experimental Crosses: Mapping Polygenic Traits, Including QTLs

Experimental crosses of mice and rats offer an ideal setting for genetic dissection of mammalian physiology (Fig. 4). With the opportunity to study hundreds of meioses from a single set of parents, the problem of genetic heterogeneity disappears, and far more complex genetic interactions can be probed than is possible for human families. Animal studies are thus an extremely powerful tool for extending the reach of genetic analysis. Of course, animal studies must always be evaluated for their applicability to the study of human diseases. Because disease-causing mutations may occur at many steps in a pathway, animal models may not point to those genes most frequently mutated in human disease. However, animal studies should identify key genes acting in the same biochemical pathway or physiological system. Animal models that are poor models for pharmacologists seeking to evaluate a new human drug therapy may nonetheless be excellent models for geneticists seeking to elucidate the possible molecular mechanisms or pathways affected in a disease.

The power of experimental crosses is most dramatically seen in the ability to dissect quantitative traits into discrete genetic factors (112). Systematic quantitative trait locus (QTL) mapping has only recently become possible with the construction of dense genetic linkage maps for mouse and rat (18, 113, 114) and the development of a suitable analytical approach for a whole-genome search, known as interval mapping. Interval mapping uses phenotypic and ge-



## ARTICLES

netic marker information to estimate the probable genotype and the most likely QTL effect at every point in the genome, by means of a maximum-likelihood linkage analysis. The basic method was introduced by Lander and Botstein for a simple situation (47) but has been generalized to a wide variety of settings (59, 115, 116). In general, QTL mapping is much more powerful in experimental crosses than in human families because of the fundamental differences in the statistical comparisons involved (117) and because nongenetic noise can be decreased through the use of progeny tests, recombinant inbred strains, and recombinant congenic strains (47, 118).

Genome-wide QTL analysis was first applied to fruit characteristics in the tomato (119), but it was soon used in mammals to study epilepsy in mice and hypertension in rats (113, 120). In the latter case, the animal study rapidly stimulated parallel human studies, with the reported linkage of the ACE gene to hypertension in rats provoking investigation of various genes in the pathway and leading to the implication of angiotensinogen in essential hypertension in humans. In only a short time, there has been an explosion of interest in QTL mapping in both agriculture and biomedicine (121). The approach opens the way to understanding the genetic basis for the tremendous strain variations seen in such quantitative traits as cancer susceptibility, drug sensitivity, resistance to infection, and aggressive behavior (122). The most important application of QTL mapping may turn out to be the identification of modifier genes affecting single-gene traits. Yeast geneticists routinely use suppressor analysis to study a mutant gene by isolating secondary mutations capable of modifying the original mutant phenotype. Although mammalian geneticists cannot easily use mutagenesis to find suppressors, they may be able to accomplish the same goal by breeding mutations onto different genetic backgrounds and dissecting the QTLs that affect the phenotypic expression. A first such example is the finding that intestinal neoplasias induced by mutations in the mouse *Apc* gene can be dramatically influenced by a modifier locus on chromosome 4 (18). By

applying this approach to the ever-growing list of gene knockouts, it should be possible to identify many additional interacting genes.

Experimental crosses also facilitate analysis of discrete traits with complex genetic etiology. Studies of type I diabetes in the nonobese diabetic mouse report the mapping of a dozen loci, each making a partial contribution to a threshold trait (123). Analysis of type I diabetes in the BB rat points to a purely synthetic interaction with one, two, or three genes required to produce disease, depending on the particular cross (124).

After initial mapping, experimental geneticists can study the physiological effects of individual polygenic factors by constructing congenic strains that differ only in the region of a single locus. Genes may also be mapped more finely by systematically whittling away at the size of the congenic interval. In some cases, synteny conservation in gene order between different mammals may point to interesting regions to investigate in the human genome.

An important point about the use of experimental crosses deserves to be emphasized, because it is commonly misunderstood. Genetic mapping results need not be consistent among different crosses. Linkage analysis reveals only those trait-causing genes that differ between the two parental strains used. A QTL may thus be detected in an  $A \times B$  cross, but not in an  $A \times C$  cross. Moreover, the effect of a QTL allele may change—or even disappear—when bred onto a different genetic background, because of epistatic effects of other genes.

### Statistical Significance

One of the thorniest problems in the genetic analysis of complex traits is to know whether a result is statistically significant. Psychiatric genetics has confronted this issue most squarely, as reported linkages to manic depression or schizophrenia have typically failed to withstand close scrutiny or replication (57, 125). Statistical significance is a challenging problem because genetic analysis can involve two types of fishing expeditions: testing many chromosomal

regions across a genome and testing multiple models for inheritance.

For example, human geneticists have long used the convention that a lod score exceeding 3 should be required to declare linkage to a simple Mendelian trait. The threshold was based on a Bayesian argument involving the prior probability of finding a gene and aimed to yield a false positive rate of 5%. Unfortunately, the reasoning does not extend well to the modern world of complex traits (with no clear prior hypothesis) or dense maps (with thousands of markers that can be tested). Instead, two approaches have gained favor in recent years.

**Analytical methods.** Formally speaking, genetic dissection involves calculating a statistic  $X$  throughout a genome. The issue of statistical significance consists of choosing an appropriate threshold  $T$  for declaring the presence of a gene, such that the genome-wide false positive rate,  $\text{Prob}(X > T)$ , is small, under the null hypothesis that no gene is present. In some cases, the genome-wide false positive rate can be estimated on the basis of simple and elegant mathematical formulas. The unifying idea comes from the insight (47, 126) that many linkage statistics tend to an asymptotic null distribution that is closely related to a well-known physical process called the Ornstein-Uhlenbeck diffusion (which describes the velocity of a particle undergoing one-dimensional Brownian motion). The problem of random large excursions of such diffusions has been extensively studied and applies directly to genetic analysis. The genome-wide false positive rate,  $\alpha_T^* = \text{Prob}(X > T \text{ somewhere in the genome})$ , can be related to the nominal false positive rate,  $\alpha_T = \text{Prob}(X > T \text{ at a single point})$ , by the formula

$$\alpha_T^* = [C + 2pGh(T)]\alpha_T$$

where  $C$  is the number of chromosomes,  $G$  is genetic length of the genome in Morgans, and the constant  $p$  and the function  $h(T)$  are defined in the notes (127). Solving  $\alpha_T^* = 0.05$  yields the appropriate threshold  $T$ . As confirmed by simulation studies, the estimates apply well to many basic situations (47, 128). Appropriate thresholds for various settings are shown in Table 1. For traditional human linkage analysis, the appropriate asymptotic lod score threshold for a 5% significance level is about 3.3. The traditional threshold of 3 actually yields a genome-wide false positive rate of about 9%. Note that all of the thresholds correspond to nominal  $p$  values less than  $10^{-4}$ ; this is considerably more stringent than the level of  $10^{-3}$  applied by many authors.

The problem of searching over alternative models has received formal attention in only a few cases (61). Current practice is to

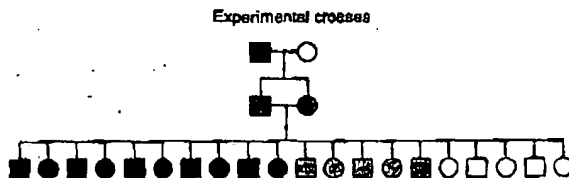


Fig. 4. Experimental crosses can provide a large number of progeny while ensuring genetic homogeneity. As a result, experimental crosses permit the genetic dissection of more complex genetic interactions than directly possible in human families, such as mapping of QTLs. Examples include epilepsy in mice, hypertension in rats, type I diabetes in mice and rats, and susceptibility to intestinal cancer in mice.



consider that each of the  $k$  models examined are statistically independent. The traditional Bonferroni correction prescribes multiplying the significance level by  $k$  or, equivalently, increasing the required lod score threshold by about  $\log_{10}(k)$  (129). The approach will likely be too conservative if the models are dependent.

**Simulation studies.** Unfortunately, the analytical approach depends on key assumptions (such as the normality of an underlying statistic and the pooling of many independent meioses) which will often be false in important situations, for example, affected-pedigree-member (APM) analysis of a modest number of large pedigrees. The best approach in such cases is to directly estimate the false positive rate by simulation. In most settings, one can randomly generate the inheritance pattern of genetic markers in a pedigree according to the laws of Mendelian inheritance and then recalculate the value of the statistic  $X$  for each such replicate (61, 130). In some settings, one can apply permutation tests such as scrambling the phenotypes or genotypes in a sib pair or QTL analysis (131). Simulation-based tests have received a great deal of attention in statistics in general (132) and are very appropriate for many genetic analyses settings (61, 130, 131). They have been applied to the

problem of genome-wide search and model selection (61). We strongly advocate this approach, although broad use will require increased dissemination of computer programs for simulation analysis.

A final issue should be noted. The appropriate thresholds for whole-genome searches should always be applied to any new hypothesis, even if one only searches over a small subset of the genome. The reason is that traits of interest will typically be studied by multiple investigators, but only positive results will be published. The genetics community as a whole is thus conducting a whole-genome scan, and the full multiple testing threshold should be applied to any positive result. Some authors have suggested avoiding this problem by developing hypotheses in one data set and retesting them in another (133). This can be helpful, but one must still apply a correction if one expects to retest multiple hypotheses at the second stage.

### Experimental Design

In designing a genetic dissection, two crucial choices arise: (i) the number and type of families from which to collect data and (ii) the number and type of genetic markers to use. To make these choices, one needs to

know the statistical power to detect a gene as a function of these choices.

For a simple Mendelian monogenic trait, a basic rule of thumb suffices: With a genetic map containing highly polymorphic markers every 20 centimorgans, linkage can be easily detected with about 40 informative meioses (21, 134). More generally, the power to detect linkage depends essentially on the number of informative meioses, almost regardless of family structure. Power can be approximated simply by counting informative meioses and can be more precisely estimated with simulation-based computer packages such as SIMLINK and SLINK (135).

In contrast, there is no comparable prescription for a complex trait. The optimal experimental design depends on the precise details of the genetic complexities, information which is typically not known in advance. The best compromise is to design a study to have sufficient power to detect any genes with effects exceeding a given magnitude. For example, one can calculate the number of sib pairs required to use allele-sharing methods to detect a locus that increases the relative risk to siblings by at least twofold (32, 82, 136). However, even if the overall relative risk to siblings is large, there is no guarantee that there exists any individual locus having an effect of this magnitude. Similarly, one can calculate the number of progeny needed to detect a QTL accounting for 10% of the phenotypic variance of a trait, but predicting whether any such loci will be present is possible only under very favorable circumstances (137). Genetic analyses of complex traits should always explicitly report the minimum effect that could have been reliably detected given the subjects studied.

The optimal choice of which families or crosses to study may also vary with the circumstances. For human studies, the range of choices include whether to focus on individuals with extreme phenotypes, when to extend a pedigree, and whether to prefer or to exclude families with too many affected individuals (137). For animal studies, the issues include whether to set up a backcross or intercross and whether to concentrate on the progeny with the most extreme phenotypes (47, 138).

The optimal density of genetic markers is a topic requiring more attention. The effect of polymorphism rate on the power of allele-sharing methods has been studied for single markers (33, 95, 136, 139), but not for the more realistic situation of multipoint mapping. It is clear that denser maps are needed for the study of sib pairs without available parents or for the study of more distant relatives, but quantitative guidance is lacking. The effect of marker density on experimental crosses has been more exten-

Table 1. Asymptotic thresholds corresponding to a genome-wide significance level of 5%. The human applications pertain to linkage analysis in a pedigree without many missing individuals and to the most common types of affected-relative pair analysis. For sib pairs, two cases are considered: (i) parents available for typing, with allele sharing on maternal and paternal chromosomes counted separately, and (ii) parents unavailable for typing, in which case sibs share zero, one, or two copies and the possible triangle method is applied (see (81)). The mouse applications pertain to situations such as QTL mapping, in which either a single parameter is estimated (for example, an additive effect) or two parameters are estimated (for example, independent additive and dominant effects). Two-sided tests are used to allow for either parental strain to contribute alleles that increase the trait. Asymptotic thresholds correspond to the situation of a dense genetic map applied to a large number of meioses. Equivalent thresholds are given in terms of lod scores, normal scores used in allele-sharing methods and QTL mapping (that is, standard random normal variables  $Z$ ) or, more generally, square-roots of  $\chi^2$  variables ( $\sqrt{\chi^2}$ ), and nominal  $p$  values for a single point. The appropriate asymptotic threshold is derived by setting  $[C + 2pG \ln(7)]/\alpha = 0.05$ ; see (127) for details. The assumed genome size is 3300 cM for the human and 1600 cM for the mouse. The slightly larger genetic length of 4000 cM for the human increases the required lod and normal-variable thresholds by about 0.1 units and decreases the  $p$  values by about 25%.

Application	Test	Lod score	Normal score	$p$ value
<b>Human</b>				
Standard linkage analysis, one free parameter	One-sided	3.3	3.9 ( $Z$ )	$5 \times 10^{-5}$
Allele sharing: Grandparent-grandchild pairs	One-sided	3.3	3.9 ( $Z$ )	$5 \times 10^{-5}$
Allele sharing: Half-sib or sib pairs (parents typed)	One-sided	3.6	4.1 ( $Z$ )	$3 \times 10^{-5}$
Allele sharing: Sib pairs (parents untyped)	Possible triangle	4.0	$4.3(\sqrt{\chi^2} 2df)$	$2 \times 10^{-5}$
Allele sharing: Uncle-nephew or first-cousin pairs	One-sided	3.7	4.1 ( $Z$ )	$2 \times 10^{-5}$
<b>Mouse</b>				
Backcross or intercross: 1 degree of freedom	Two-sided	3.3	3.9 ( $Z$ )	$1 \times 10^{-4}$
Intercross: 2 degrees of freedom	Two-sided	4.3	$4.4(\sqrt{\chi^2} 2df)$	$5 \times 10^{-5}$
Recombinant inbred lines: 1 degree of freedom	Two-sided	3.9	4.3 ( $Z$ )	$2 \times 10^{-5}$

## ARTICLES

sively studied (47, 140). Finally, a few authors have begun to explore two-tiered strategies, in which initial evidence is obtained with a sparse map and then confirmed with a dense map (141).

### Cloning Genes That Underlie Complex Traits

Once genetic dissection implicates a chromosomal region, there remains the formidable task of identifying the responsible gene. That type 1 diabetes cosegregates with anonymous markers on chromosome 11q in the human or that hypertension cosegregates with the ACE gene in rat crosses simply indicates that a causative gene lies somewhere nearby. However, the possible region might be as large as 10 to 20 Mb—enough to contain 500 genes. Positional cloning requires higher resolution mapping to narrow the search to a tractable region.

For a simple Mendelian trait, the situation is most favorable. Because the responsible gene must show perfect cosegregation with the trait, even a single crossover suffices to eliminate a region from consideration. From a study of 200 meioses, the interval can be pared to about 1 cM, corresponding to about 1 Mb (142). Still, the challenge is considerable. It is sobering to note that virtually all successful positional cloning efforts have depended on the fortuitous presence of chromosomal aberrations, trinucleotide repeat expansions, or previously known candidate genes. Only two human disease genes have been positionally cloned solely on the basis of point mutations: cystic fibrosis and diastrophic dysplasia (DTD) (143).

For complex traits, positional cloning will likely be even harder. Because cosegregation is not expected to be perfect, single crossovers no longer suffice for fine-structure mapping. Resolution becomes a statistical matter (144). For a gene conferring a relative risk of twofold, for example, one would need to examine a median number of nearly 600 sib pairs to narrow the likely region (95% confidence interval) to 1 cM. Moreover, the genes underlying complex traits may be subtle missense mutations rather than gross deletions. How will positional cloners overcome these obstacles?

In the human, the most powerful strategy may prove to be linkage disequilibrium mapping in genetically isolated populations (21, 145). The idea is to find many affected individuals who have inherited the same disease-causing allele from a common ancestor. Such individuals will tend to have retained the particular pattern of alleles present on the ancestral chromosome, with the immediate vicinity of the gene being evident as the region of maximal retention. In effect, the method exploits information

from many historical meioses and thereby affords much higher recombinational resolution. Fine-structure linkage disequilibrium mapping has been applied to the isolated Finnish population (founded about 100 generations ago) to permit the cloning of the DTD gene (143). Whereas conventional recombinational mapping was only able to localize the gene to within about 1.5 cM, linkage disequilibrium studies were able to pinpoint it to within about 50 kb. The approach is also applicable to younger populations: linkage disequilibrium should be detectable over larger distances, although the ultimate resolving power will be less (146). Elegant studies in the Mennonite population (founded about 10 generations ago) have allowed initial mapping of genes involved in a recessive form of Hirschsprung disease (20).

In animal models, fine-structure mapping of factors such as QTLs can be accomplished through appropriate breeding. The key is to ensure unambiguous genotyping at the trait-causing locus. The best solution is probably to (i) create congenic strains differing only in the region of interest, (ii) cross these strains to construct recombinant chromosomes (that is, ones in which there has been a crossover between flanking genetic markers), and (iii) evaluate each recombinant chromosome to determine which trait-causing allele is carried by performing progeny testing (that is, examining the phenotype of many progeny carrying the chromosome) (113). The construction of the required congenic strains would traditionally require 20 generations of breeding. With the advent of complete genetic linkage maps, however, one can construct "speed congenics" in only three to four generations by using marker-directed breeding (147).

The Human Genome Project promises to make a tremendous contribution to the positional cloning of complex traits by eventually providing a complete catalog of all genes in a relevant region. With such information, positional cloning will be reduced to the systematic evaluation of candidate genes—still challenging, but far more manageable than today's more haphazard forays. Indeed, the Human Genome Project is essential if the genetic analysis of complex traits is to achieve its full potential.

Finally, candidate genes, whether identified by positional cloning or guessed a priori, must always be subjected to rigorous evaluation before they are accepted. The gold-standard tests for human genes should include association studies demonstrating a clear correlation between functionally relevant allelic variations and the risk of disease in humans, and transgenic studies demonstrating that gene addition or gene knock-out in animals produces a phenotypic effect.

For genes identified from experimental animal crosses, one can and should go a step further by demonstrating that an induced knockout allele at the candidate gene fails to complement an allele at the locus to be cloned (148).

### Conclusion

In the early 1900s, the fledgling theory of Mendelian genetics was attacked on the grounds that the simple, discrete inheritance patterns of pea shape or *Drosophila* eye color did not apply to the variation typically seen in nature (149). After 20 years of acrimonious battle, the issue was eventually resolved with the theoretical understanding that Mendelian factors could give rise to complex and continuous traits, even if direct identification of the genes themselves was not practical. Now, with the advent of dense genetic linkage maps, geneticists are taking up the challenge of the genetic dissection of complex traits. If they are successful, the tools of genetics will be brought to bear on some of the most important problems in human health and in agriculture, and the Mendelian revolution will finally be complete.

### REFERENCES AND NOTES

1. A. H. Sturtevant, *J. Exp. Zool.* 14, 43 (1913).
2. W. Bender, P. Spierer, D. S. Hogness, *J. Mol. Biol.* 168, 17 (1983); P. Spierer, A. Spierer, W. Bender, D. S. Hogness, *ibid.*, p. 35.
3. T. D. Petes and D. Botstein, *Proc. Natl. Acad. Sci. U.S.A.* 74, 5091 (1977).
4. D. Botstein, R. L. White, M. H. Skolnick, R. W. Davis, *Am. J. Hum. Genet.* 32, 314 (1980).
5. D. N. Cooper and J. Schmidtke, *Ann. Med.* 24, 29 (1992); V. A. McKusick, *Mendelian Inheritance in Man: Catalogs of Autosomal Dominant, Autosomal Recessive, and X-Linked Phenotypes* (Johns Hopkins Univ. Press, Baltimore, MD, 1992); F. Collins, personal communication.
6. T. H. J. Huisman, *Am. J. Hum. Genet.* 173 (1979); M. H. Steinberg and R. P. Hebbel, *ibid.* 14, 405 (1983).
7. G. J. Dover et al., *Blood* 80, 818 (1992); S. L. Thiel et al., *Am. J. Hum. Genet.* 54, 214 (1994).
8. D. Easton, D. Ford, J. Peto, *Cancer Surv.* 18, 95 (1993); D. Ford et al., *Lancet* 343, 692 (1994).
9. S. T. Heiders et al., *Hum. Genet.* 76, 348 (1987); G. Romeo et al., *Lancet* II, 8 (1988); A. Turco, B. Paisiol, L. Gammara, G. Maschio, P. F. Pignatti, *Clin. Genet.* 40, 287 (1991).
10. P. H. St. George-Hyslop et al., *Nature* 347, 194 (1990).
11. J. Barbosa, H. Noren, F. C. Goutz, E. J. Yunis, *Diabetes Metab.* 2, 100 (1976); G. I. Bell et al., *Proc. Natl. Acad. Sci. U.S.A.* 88, 1484 (1991); D. W. Bowden et al., *Am. J. Hum. Genet.* 50, 607 (1992); P. Froguel et al., *Nature* 356, 162 (1992).
12. R. Fishel et al., *Cell* 78, 1027 (1993); F. Leach et al., *ibid.*, p. 1215; N. Papadopoulos et al., *Science* 263, 1625 (1994); R. Fishel et al., *Cell* 77, 187 (1994); B. Liu et al., *Cancer Res.* 54, 4590 (1994).
13. N. G. Jaspers and D. Bootsma, *Proc. Natl. Acad. Sci. U.S.A.* 79, 2641 (1982); E. Sobel et al., *Am. J. Hum. Genet.* 50, 1343 (1992).
14. E. A. De Werd-Kastelen, W. Kalfer, D. Bootsma, *Nat. New Biol.* 238, 80 (1972).
15. L. M. Bleeker-Wagemakers et al., *Genomics* 14, 811 (1992); R. Kumar-Singh, P. F. Kenna, G. J. Farrar, P. Humphries, *ibid.* 16, 212 (1993).
16. S. Bril et al., *J. Clin. Invest.* 81, 1710 (1988); J. M. Tager et al., *Prog. Clin. Biol. Res.* 321, 545 (1990).

- J. Gartner, H. Moser, D. Valle, *Nat. Genet.* 1, 16 (1992); N. Shimozono et al., *Science* 256, 1132 (1992).
17. R. Ward, in *Hypertension: Pathophysiology, Diagnosis, and Management*, J. H. Laragh and B. M. Brenner, Eds. (Raven, New York, 1990), pp. 81-100.
  18. W. F. Dietrich et al., *Cell* 75, 631 (1993).
  19. K. Kajiwara, E. L. Berson, T. P. Dryja, *Science* 264, 1604 (1994).
  20. E. G. Puffenberger et al., *Hum. Mol. Genet.* 3, 1217 (1994).
  21. E. S. Lander and D. Botstein, *Cold Spring Harbor Symp. Quant. Biol.* 61, 49 (1996).
  22. C. C. Li, *Am. J. Hum. Genet.* 41, 517 (1987); P. P. Majumder, S. K. Das, C. C. Li, *ibid.* 43, 119 (1988); S. K. Nath, P. P. Majumder, J. J. Nordlund, *ibid.*, in press.
  23. Bilineality has been thought to plague many genetic analyses, particularly those involving psychiatric disorders. See, for example, S. E. Hodges, *Genet. Epidemiol.* 9, 181 (1992); M. Dumer, D. A. Greenberg, S. E. Hodges, *Am. J. Hum. Genet.* 51, 859 (1992); K. R. Markianos, M. A. Spence, D. J. Kupfer, *Arch. Gen. Psychiatry* 48, 1137 (1991); M. A. Spence et al., *Hum. Hered.* 43, 166 (1993).
  24. A simple application of the Hardy-Weinberg principle shows that the proportion of individuals who inherit D from a homozygous parent (as opposed to a heterozygous parent) is equal to the allele frequency of D. Suppose that allele D has frequency  $p$  and let  $q = 1 - p$ . The proportion of individuals in the population who inherit D from a homozygous parent is  $2p^2/(2p^2 + 2pq) = p$ .
  25. M. A. Pericak-Vance et al., *Am. J. Hum. Genet.* 48, 1034 (1991).
  26. E. H. Cordier et al., *Science* 261, 921 (1993).
  27. D. C. Wallace, *Annu. Rev. Biochem.* 61, 1175 (1992); J. B. Redman, R. G. Fenwick Jr., Y. H. Fu, A. Pizzulli, C. T. Caskey, *J. Am. Med. Assoc.* 269, 1960 (1993).
  28. N. J. Schork and S. W. Guo, *Am. J. Hum. Genet.* 53, 1320 (1993).
  29. M. J. Khoury, T. H. Beaty, B. H. Cohen, *Fundamentals of Genetic Epidemiology* (Oxford Univ. Press, New York, 1993); K. M. Weiss, *Genetic Variation and Human Disease* (Cambridge Univ. Press, Cambridge, 1993).
  30. M. C. Neale and L. R. Cardon, *Methodology for Genetic Studies of Twins and Families* (Kluwer Academic, Boston, 1992).
  31. L. S. Penrose, *Acta Genet.* 4, 257 (1953); J. H. Edwards, *ibid.* 10, 63 (1960); *Br. Med. Bull.* 25, 58 (1969); N. Risch, *Am. J. Hum. Genet.* 46, 222 (1990).
  32. N. Risch, *Am. J. Hum. Genet.* 46, 229 (1990).
  33. ———, *ibid.*, p. 242.
  34. The concept of heritability has a long history in genetics, being one of the most widely used terms in all of biology; see, for example, R. A. Fisher, *Trans. R. Soc. Edinburgh* 82, 399 (1918); D. S. Falconer, *Introduction to Quantitative Genetics* (Longman, New York, 1981).
  35. B. Newman, M. A. Austin, M. Lee, M. C. King, *Proc. Natl. Acad. Sci. U.S.A.* 85, 3044 (1988).
  36. See also W. R. Williams and D. E. Anderson, *Genet. Epidemiol.* 1, 7 (1984); D. T. Bishop, L. Cannon-Albright, T. McEellan, E. J. Gardner, M. H. Skolnick, *Genet. Epidemiol.* 5, 151 (1988); E. B. Claus, N. J. Risch, W. D. Thompson, *Am. J. Epidemiol.* 131, 881 (1990).
  37. Problems that arise through the ascertainment of probands whose families are then used in genetic studies have been discussed at length in the genetics literature. Some articles of relevance are W. Weinberg, *Arch. Rassen-Gesellschaftsbiol.* 9, 165 (1912); N. E. Morton, *Am. J. Hum. Genet.* 11, 1 (1959); W. J. Ewens, *ibid.* 34, 853 (1982); P. P. Majumder, *Stat. Med.* 4, 163 (1985); W. J. Ewens and N. C. Shute, *Ann. Hum. Genet.* 50, 399 (1986); M. R. Young, M. Boehnke, P. P. Mol, *Am. J. Hum. Genet.* 43, 705 (1988).
  38. J. Ott, *Am. J. Hum. Genet.* 48, 219 (1990).
  39. The number of major genetic factors segregating in a cross between inbred strains can be estimated under some circumstances; see (47); W. E. Castle, *Science* 64, 223 (1916); S. Wright, *Evolution and the Genetics of Populations: Genetic and Biometric Foundations* (Univ. of Chicago Press, Chicago, 1968), pp. 373-420.
  40. M. W. Feldman and L. L. Cavalli-Sforza, in *Genetic Aspects of Common Diseases: Applications to Predictive Factors in Coronary Disease* (Liss, New York, 1979), pp. 203-227.
  41. K. W. Kinzler et al., *Science* 253, 661 (1991); G. Joslyn et al., *Cell* 66, 601 (1991).
  42. L. A. Aittonian et al., *Science* 260, 612 (1993).
  43. R. R. Williams et al., *Arch. Intern. Med.* 150, 582 (1990); J. V. Selby et al., *J. Am. Med. Assoc.* 265, 9 (1991); R. R. Williams et al., *Ann. Med.* 24, 469 (1992).
  44. R. F. Clark and A. M. Goate, *Arch. Neurol.* 50, 1164 (1993).
  45. M. E. Maronberg, N. Risch, L. F. Berkman, B. Flodorus, U. de Faire, *N. Engl. J. Med.* 330, 1041 (1994).
  46. G. Carey and J. Williamson, *Am. J. Hum. Genet.* 48, 788 (1991); L. R. Cardon and D. W. Fulker, *Am. J. Hum. Genet.*, in press; N. J. Schork et al., in preparation.
  47. E. S. Lander and D. Botstein, *Genetics* 121, 174 (1989).
  48. L. Groop and R. DeFranco, personal communication.
  49. G. A. Barnard, *R. Stat. Soc. J. B11*, 115 (1949); J. B. S. Haldane and C. A. B. Smith, *Ann. Eugen.* 14, 10 (1947); J. Chotot, *Ann. Hum. Genet.* 48, 359 (1984).
  50. N. E. Morton, *Am. J. Hum. Genet.* 8, 80 (1955).
  51. A. W. F. Edwards, *Likelihood* (Johns Hopkins Univ. Press, Baltimore, MD, 1982).
  52. J. Ott, *Analysis of Human Genetic Linkage* (Johns Hopkins Univ. Press, Baltimore, MD, 1991), pp. 65-68.
  53. J. Tomfohrde et al., *Science* 264, 1141 (1994).
  54. J. K. Ghosh and P. K. Sen, in *The Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer*, L. M. LeCam and R. A. Olshen, Eds. (Wadsworth, Monterey, CA, 1985), pp. 807-810; D. M. Titterton, A. F. M. Smith, U. E. Makov, *Statistical Analysis of Finite Mixture Distributions* (Wiley, London, 1985); J. Ott, in (52), pp. 194-216; J. Faraway, *Genet. Epidemiol.* 10, 75 (1993); H. Chernoff and E. S. Lander, *J. Stat. Plann. Inference*, in press.
  55. J. M. Hall et al., *Science* 260, 1684 (1990).
  56. G. D. Schellenberg et al., *ibid.* 258, 688 (1992); C. E. Yu et al., *Am. J. Hum. Genet.* 54, 631 (1994).
  57. R. Sherrington et al., *Nature* 338, 164 (1988).
  58. E. S. Lander and D. Botstein, *Proc. Natl. Acad. Sci. U.S.A.* 85, 7353 (1988); N. J. Schork, M. Boehnke, J. D. Terwilliger, *J. Ott, Am. J. Hum. Genet.* 53, 1127 (1993).
  59. C. S. Haley and S. A. Knott, *Heredity* 68, 315 (1992).
  60. P. J. Triemer et al., *Genomics* 19, 320 (1994).
  61. D. E. Weeks, T. Lohner, E. Squires-Wheeler, C. Kaufmann, *J. Ott, Genet. Epidemiol.* 7, 237 (1990).
  62. One way of avoiding the possible use of an incorrectly specified segregation model in a linkage analysis is to estimate linkage and segregation parameters simultaneously. Although computationally prohibitive in certain settings, this strategy appears to have some favorable features. See, for example, C. J. MacLean, N. E. Morton, S. Yee, *Comput. Biomed. Res.* 17, 471 (1984); G. E. Bonney, G. M. Lathrop, J. M. Lalouel, *Am. J. Hum. Genet.* 43, 29 (1988); N. J. Schork, *Genet. Epidemiol.* 9, 207 (1992).
  63. J. Ott, *Clin. Genet.* 12, 119 (1977); F. Clerget-Darpoux, *Ann. Hum. Genet.* 46, 363 (1982); ———, C. Bonati-Pellie, J. Hochoz, *Biometrika* 42, 393 (1986); F. Clerget-Darpoux, M. C. Babron, C. Bonati-Pellie, *J. Psychiatr. Res.* 21, 825 (1987).
  64. For a discussion of how robust linkage analysis might be in certain settings, see, for example, J. A. Williamson and C. I. Amos, *Genet. Epidemiol.* 7, 309 (1990); C. I. Amos and J. A. Williamson, *Am. J. Hum. Genet.* 52, 213 (1993); S. E. Hodges and R. C. Elston, *Genet. Epidemiol.*, in press.
  65. L. J. Eaves, *Heredity* 72, 175 (1994).
  66. J. Ott, *Am. J. Hum. Genet.* 31, 161 (1979); J. Edwards, *Cytogenet. Cell Genet.* 39, 43 (1982); N. J. Schork, in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, Seattle, WA, 21-24 April 1994, E. M. Koramides and S. M. Kaufman, Eds. (Interface Foundation of North America, Fairfax Station, VA, 1992); T. J. Goradia, K. Lange, P. L. Miller, P. M. Mackinnon, *Hum. Hered.* 42, 42 (1992); R. W. Cottingham Jr., R. M. Idury, A. A. Schaffer, *Am. J. Hum. Genet.* 53, 252 (1993).
  67. N. J. Schork, *Genet. Epidemiol.* 8, 29 (1991).
  68. R. C. Elston and J. Stewart, *Hum. Hered.* 21, 20 (1971).
  69. Since its introduction, the Elston-Stewart algorithm has been extended and modified in various ways; see, for example, J. Ott, *Am. J. Hum. Genet.* 26, 588 (1974); K. Lange and R. C. Elston, *Hum. Hered.* 26, 95 (1975); C. Cannings, E. T. Thompson, M. H. Skolnick, *Adv. Appl. Probab.* 10, 20 (1976); R. C. Elston, V. T. George, F. Severnson, *Hum. Hered.* 42, 16 (1992).
  70. J. Ott, *Am. J. Hum. Genet.* 28, 526 (1978).
  71. G. M. Lathrop, J. M. Lalouel, C. Juller, *J. Ott, Proc. Natl. Acad. Sci. U.S.A.* 81, 3443 (1984).
  72. N. E. Morton and C. J. MacLean, *Am. J. Hum. Genet.* 28, 469 (1974); S. J. Hassstedt, *Comput. Biomed. Res.* 15, 295 (1982); G. E. Bonney, *Am. J. Med. Genet.* 18, 731 (1984); S. W. Guo and E. T. Thompson, *IMA J. Math. Appl. Med. Biol.* 8, 171 (1991); *ibid.*, p. 149; N. J. Schork, *Genet. Epidemiol.* 8, 73 (1992).
  73. M. C. King, G. M. Lee, N. B. Spinner, G. Thompson, M. R. Wrensch, *Annu. Rev. Public Health* 6, 1 (1984); E. A. Thompson, *Stat. Med.* 5, 291 (1986).
  74. E. S. Lander and P. Green, *Proc. Natl. Acad. Sci. U.S.A.* 84, 2363 (1987).
  75. L. Kruglyak and E. S. Lander, in preparation.
  76. S. W. Guo and E. T. Thompson, *Am. J. Hum. Genet.* 51, 1111 (1992); D. C. Thomas, *Cytogenet. Cell Genet.* 58, 228 (1992); A. Kong, N. Cox, M. Flogge, M. Irwin, *Genet. Epidemiol.* 10, 483 (1993).
  77. For a discussion, see W. C. Blackwelder and R. C. Elston, *Genet. Epidemiol.* 2, 85 (1985); A. S. Whittemore and J. Halpern, *Biometrics* 50, 118 (1994).
  78. D. E. Weeks and K. Lange, *Am. J. Hum. Genet.* 42, 315 (1988).
  79. B. K. Suarez, J. Rice, T. Reich, *Ann. Hum. Genet.* 42, 67 (1978).
  80. L. R. Weitkamp, *N. Engl. J. Med.* 305, 1301 (1981); M. Knapp, S. A. Beuchter, M. P. Bair, *Hum. Hered.* 44, 97 (1994).
  81. P. Holmans, *Am. J. Hum. Genet.* 52, 362 (1993).
  82. N. E. Morton et al., *ibid.* 35, 201 (1983); R. N. Hyer et al., *ibid.* 48, 243 (1991); S. S. Rich, S. S. Parmer, F. C. Goetz, B. Hedlund, J. Barbosa, *Diabetologia* 34, 350 (1991).
  83. D. Owerbach and K. H. Gabbay, *Am. J. Hum. Genet.* 54, 909 (1994).
  84. J. Davies et al., *Nature*, in press.
  85. D. H. Hamer, S. Hu, V. L. Magnuson, N. Hu, A. M. L. Pettaucci, *Science* 261, 321 (1993).
  86. In principle, one could either explicitly use the expected value of the IBD statistic or work with the probability distribution over the IBD statistic. The former approach has the difficulty that it is hard to derive an expression for the variance of the statistic. It has been done in some simple cases, such as inbred crosses (L. Kruglyak and E. S. Lander, submitted). The latter has been applied in numerous cases (for example, (79); C. I. Amos, D. V. Dawson, R. C. Elston, *Am. J. Hum. Genet.* 47, 842 (1990)).
  87. K. Lange, *Am. J. Hum. Genet.* 39, 148 (1986); D. E. Weeks and K. Lange, *ibid.* 50, 659 (1992).
  88. X. Jeunemaitre et al., *Cell* 71, 169 (1992).
  89. J. K. Haseman and R. C. Elston, *Behav. Genet.* 2, 3 (1972).
  90. W. C. Blackwelder and R. C. Elston, *Commun. Stat. Theor. Methods* 11, 449 (1982); C. I. Amos and R. C. Elston, *Genet. Epidemiol.* 6, 349 (1989); J. M. Olson and E. M. Wilman, *ibid.* 10, 87 (1993).
  91. C. I. Amos et al., *Genet. Epidemiol.* 3, 255 (1988).
  92. D. G. Marsh et al., *Science* 264, 1152 (1994).
  93. N. A. Morrison et al., *Nature* 367, 284 (1994).
  94. D. E. Goldgar, *Am. J. Hum. Genet.* 47, 857 (1990); C. I. Amos, *ibid.* 54, 535 (1994); D. W. Fulker and L. R. Cardon, *ibid.*, p. 1092; N. J. Schork and S. Ghosh, submitted.

## ARTICLES

95. W. H. Bertrini et al., *Proc. Natl. Acad. Sci. U.S.A.* 91, 5918 (1994).
96. L. P. Ryder, E. Andersen, A. Svegaard, Eds., *HLA and Disease Registry, Third Report* (Munksgaard, Copenhagen, 1979).
97. W. E. Braun, *HLA and Disease* (CRC, Boca Raton, FL, 1979).
98. E. H. Corder et al., *Nat. Genet.* 7, 180 (1994); C. Singer et al., *Arch. Neurol.* 49, 13 (1992); F. Schachter et al., *Nat. Genet.* 8, 29 (1994); H. Schunkert et al., *N. Engl. J. Med.* 330, 1834 (1994); F. Cambien et al., *Nature* 359, 641 (1992).
99. Some of these methods have generally failed under the heading of "measured genotype" analyses. Relevant articles include E. Boerwinkle, R. Chakraborty, C. F. Sing, *Ann. Hum. Genet.* 50, 181 (1986); V. T. George and R. C. Elston, *Genet. Epidemiol.* 4, 193 (1987).
100. An exception would be if there were so many disease-causing alleles that the association could not be easily detected, or if the trait were predominantly due to a different cause in some populations.
101. This could happen if the ancestral disease-causing chromosome carried both the Eco RI and the Bam HI sites, but many normal chromosomes also had the Eco RI site but not the Bam HI site.
102. K. M. Weiss, *Genetic Variation and Human Disease* (Cambridge Univ. Press, Cambridge, 1993).
103. C. T. Falk and P. Rubinstein, *Ann. Hum. Genet.* 51, 227 (1987); F. Clerget-Daroux et al., *ibid.* 52, 247 (1988); G. Thompson et al., *Genet. Epidemiol.* 6, 43 (1989); J. Ott, *ibid.*, p. 127; J. D. Terwilliger and J. Ott, *Hum. Hered.* 42, 337 (1992); M. Knapp, S. A. Souchter, M. P. Baur, *Am. J. Hum. Genet.* 52, 1085 (1993).
104. C. J. J. J. et al., *Nature* 354, 155 (1991).
105. R. S. S. et al., *Am. J. Hum. Genet.* 52, 506 (1993); S. E. Hodge, *ibid.* 53, 367 (1993).
106. K. Blum et al., *J. Med. Assoc.* 263, 2055 (1990).
107. J. Gelernter et al., *ibid.* 266, 1801 (1991); K. Blum et al., *Alcohol* 10, 59 (1993); J. Gelernter et al., *J. Am. Med. Assoc.* 269, 1673 (1993).
108. D. Goldman et al., *Alcohol Clin. Exp. Res.* 17, 199 (1993); B. K. Suarez et al., *Genomics* 18, 12 (1994); C. N. Pato et al., *Am. J. Med. Genet.* 48, 78 (1993); A. Parsian et al., *Arch. Gen. Psychiatry* 48, 655 (1991); T. Arimura et al., *Biol. Psychiatry* 33, 108 (1993); D. Goldman et al., *Acta Psychiatr. Scand.* 86, 351 (1992); S. Schwab et al., *Am. J. Hum. Genet.* 49 (suppl.), 203 (1991); D. Goldman et al., *Alcohol Clin. Exp. Res.*, in press.
109. Indeed, a perfect association between two loci linked at 1% recombination decays with a "half-life" of about 70 generations, because  $(0.99)^{70} \approx 1/2$ . This corresponds to about 1600 years for humans.
110. R. Plomin et al., *Behav. Genet.* 24, 107 (1994).
111. R. Chakraborty and K. M. Weiss, *Proc. Natl. Acad. Sci. U.S.A.* 85, 9118 (1988).
112. QTL analysis is included under the heading of polygenic trait analysis, despite the fact that a quantitative trait could, in principle, be monogenic. In practice, quantitative traits are nearly always polygenic, and the methodology is the same in any case.
113. H. Jacob, *Cell* 67, 213 (1991).
114. T. Senikawa et al., *Genetics* 131, 701 (1992).
115. E. A. Carbonell, T. M. Garig, E. Belsarsky, M. J. Asins, *Biometrics* 48, 305 (1992); Z. W. Luo and M. J. Kearney, *Hereditas* 89, 235 (1992); Z. B. Zeng, *Genetics* 136, 1457 (1994); S. A. Knott and C. S. Haley, *ibid.* 132, 1211 (1992); R. C. Jansen and P. Stam, *ibid.* 138, 1447 (1994); Z. B. Zeng, *Proc. Natl. Acad. Sci. U.S.A.* 90, 10872 (1993); R. C. Jansen, *Genetics* 136, 205 (1993); *Theor. Appl. Genet.* 85, 252 (1992).
116. Work on gene-mapping schemes in experimental organisms is ongoing and a topic of great interest in statistical genetics. For example, L. Kruglyak and E. S. Lander (submitted) have developed a nonparametric method for QTL mapping; N. J. Schork (submitted) has shown how to use interval mapping for dichotomous traits; and N. J. Schork (submitted) has described methods for detecting epistatic interaction.
117. The key reason is that the meioses in experimental crosses are all equivalent (that is, they involve segregation of the same alleles in a fixed phase relationship, with each parent being heterozygous for a "high" and "low" allele at a QTL; for example,  $F_2$  progeny), whereas the meioses in different human families are not directly comparable (that is, the configuration of "high" and "low" alleles varies among each set of parents, and only a subset of parents segregate for different alleles at a QTL). Accordingly, QTL analysis in experimental crosses allows one to compare means (for example, the mean phenotype of progeny inheriting allele A versus the mean phenotype of progeny inheriting allele B), whereas QTL analysis of human families requires one to compare variances (for example, squared differences  $\Delta^2$  between sib pairs inheriting  $\rightarrow$  180 alleles versus between sib pairs inheriting no alleles 180, as in the Haseman-Elston method).
118. P. C. Groer et al., *FASEB J.* 8, 2826 (1992); C. J. Moen, M. Shook, A. A. Han, P. Dement, *Oncogene* 7, 563 (1992).
119. A. H. Peterson et al., *Nature* 335, 721 (1988).
120. M. L. Rise, W. N. Frankel, J. M. Coffin, T. N. Seyfried, *Science* 253, 669 (1991); P. Hilbert et al., *Nature* 353, 521 (1991).
121. S. D. Tanksley, *Annu. Rev. Genet.* 27, 265 (1993); J. C. Crabbe, J. K. Beckman, K. J. Buck, *Science* 264, 1715 (1994).
122. M. Festing, *Inbred Strains in Biomedical Research* (Oxford Univ. Press, New York, 1979).
123. S. Ghosh et al., *Nat. Genet.* 4, 404 (1993); N. Risch, S. Ghosh, J. A. Todd, *Am. J. Hum. Genet.* 53, 702 (1993).
124. H. J. Jacob et al., *Nat. Genet.* 2, 56 (1992).
125. J. A. Egeland et al., *Nature* 328, 783 (1987); J. L. Kennedy et al., *ibid.* 336, 167 (1988); S. D. Debora-Wadleigh et al., *ibid.* 340, 391 (1989); J. R. Kesteven et al., *ibid.* 342, 238 (1989); M. Baron et al., *Nat. Genet.* 3, 49 (1993).
126. E. Feingold, P. O. Brown, D. Siegmund, *Am. J. Hum. Genet.* 53, 234 (1993).
127. The factor  $[C + 2pGh(T)]$  approximately converts a single-point significance level  $\alpha_p$  to a genome-wide significance level  $\alpha_g$ , both for lod scores and allele-sharing statistics. The factor is asymptotically correct for large values of  $T$  and thus small values of  $\alpha_p$  and  $\alpha_g$ . The significance level  $\alpha_p$  can also be regarded as the expected number of distinct regions in which the statistic  $X$  exceeds  $T$ . The significance level can be one-sided or two-sided. For a normally distributed statistic,  $h(T) = T^2$ . For a lod score statistic,  $h(T) = T/2 \log_{10} 10$ . The constant  $p$  measures the crossover rate. If  $R(c)$  denotes the covariance function between genotypes (or allele-sharing status) at distance  $c$  (measured in Morgans), then  $p$  is defined by the equation  $p = -R'(0)/2$ , where  $R'(0)$  denotes the derivative at zero taken as a limit from above. In the simple case of individual informative meioses in a nuclear family or a single-generation cross,  $R(c) = \exp(-2c)$ . More generally,  $R(c)$  is of the form  $a \exp(-b_1 c) + a_2 \exp(-b_2 c) + \dots + a_k \exp(-b_k c)$ . For example, first-cousin allele sharing has  $R(c) = \exp(-4c/2) + \exp(-8c/3) + \exp(-8c/6)$ . One can show that  $p = 1$  for backcrosses or intercrosses in which only an additive component is fit (1 df);  $p = 1.5$  for intercrosses in which an additive and dominance component is fit (2 df);  $p = 1$  for grandparent-grandchild pair allele sharing;  $p = 2$  for half-sib pair allele sharing (as well as for sib pairs regarded as the sum of two half-sib pairs) and for recombinant inbred lines produced by selfing;  $p = 2.5$  for uncle-nephew pair allele sharing;  $p = 8/3$  for first-cousin pair allele sharing; and  $p = 4$  for recombinant inbred strains produced by brother-sister mating. The analysis has also been modified by D. Siegmund (personal communication) for sib pair allele sharing in which the significance test is based on the possible triangle method (87). The calculated values in Table 1 are from L. Kruglyak, D. Siegmund, E. S. Lander, in preparation.
128. A. Reibel, B. Goltner, B. Mangin, *Genetics* 138, 235 (1994); W. N. Frankel et al., *ibid.*, in press.
129. K. K. Kidd and J. Ott, *Cytogenet. Cell Genet.* 37, 510 (1984).
130. J. Ott, *Proc. Natl. Acad. Sci. U.S.A.* 86, 4175 (1989); J. D. Terwilliger and J. Ott, *Cytogenet. Cell Genet.* 59, 142 (1992); M. N. Okano and J. R. W. Yates, *Ann. Hum. Genet.* 58, 129 (1994); N. J. Schork, in *Exploring the Limits of the Bootstrap*, R. LaPage and L. Billard, Eds. (Wiley, New York, 1991); ——— and M. A. Schork, *Am. J. Hum. Genet.* 45, 803 (1989); A. Kong, M. Frigge, M. Irwin, N. Cox, *ibid.* 61, 1413 (1992).
131. To illustrate the flexibility of permutation tests, suppose that a genome scan detects two loci on chromosomes 1 and 2 that appear to affect a trait in an additive fashion, with the first explaining twice as much variance as the second. To determine whether the second locus makes a significant contribution over and above the first, one could permute the genotypes for chromosome 2 among the progeny. The likelihood ratio statistic for a two-locus effect can be computed for each such permutation and the empirical distribution used to make inferences about the significance of the effect of the second locus.
132. See, for example, E. S. Edgington, *Randomization Tests* (Dekker, New York, 1987); E. W. Norman, *Computer Intensive Methods for Testing Hypotheses* (Wiley, New York, 1989); B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman & Hall, New York, 1994), pp. 220–226.
133. J. K. Belknap, *Behav. Genet.* 22, 677 (1992).
134. Informative meioses are those in which one can definitively tell if the disease-causing allele has been transmitted from an obligate heterozygote. For a fully penetrant monogenic dominant trait, all meioses occurring in an affected individual are informative. For a fully penetrant monogenic recessive trait, meioses occurring in an obligate carrier are informative if they produce an affected child or if they produce an unaffected child who inherits the disease-causing allele from the other parent. This latter situation occurs in one-third of unaffected children and can be recognized by the use of close flanking markers.
135. M. Boehnke, *Am. J. Hum. Genet.* 39, 613 (1986); D. E. Weeks, J. Ott, G. M. Lathrop, *ibid.* 47, A204 (1990).
136. D. Brown, M. B. Gorin, D. E. Weeks, *ibid.* 54, 544 (1994).
137. C. Cannings and E. T. Thompson, *Can. Genet.* 12, 208 (1977); M. Boehnke, M. R. Young, P. P. Moll, *Am. J. Hum. Genet.* 43, 336 (1988); A. E. Pulver et al., *Am. J. Med. Genet.* 54, 36 (1994); E. A. Thompson, *Theor. Popul. Biol.* 24, 39 (1983); *Am. J. Hum. Genet.* 39, 968 (1981); *Biometrics* 37, 313 (1981); M. Boehnke, K. H. Orntoft, J. M. Arduino, *Am. J. Hum. Genet.* 48, 581 (1990). In cases involving a disease-causing allele  $D$  with high frequency, one may also wish to avoid families with "too many" affected individuals, because there is a high probability that one or both parents is homozygous for  $D$  (L. Kruglyak and E. S. Lander, in preparation).
138. J. S. Beckman, in *Gene Mapping Techniques and Applications*, L. B. Schook, M. A. Lewin, D. G. McLaren, Eds. (Dekker, New York, 1991), pp. 201–229.
139. D. T. Bishop and J. A. Williamson, *Am. J. Hum. Genet.* 46, 254 (1990).
140. A. Darsvasi, A. Weinreb, V. Minko, J. L. Waller, M. Soller, *Genetics* 134, 943 (1993); Z. B. Zeng, *ibid.* 136, 1457 (1994).
141. R. C. Elston, in *Proceedings of the 16th International Biometric Conference*, Hamilton, New Zealand, 7–11 December 1992 (Ruakura Conference Center, Hamilton, New Zealand, 1992), pp. 39–61.
142. Crossovers are randomly distributed with density  $N$  per centimorgan, where  $N$  is the number of informative meioses. It follows trivially that the expected distance to the nearest flanking crossover on either side is  $(1/N)$  cM. [See also, K. Lange, L. Kunkel, J. Aldridge, S. A. Latt, *Am. J. Hum. Genet.* 37, 653 (1985); M. Boehnke, *ibid.* 55, 379 (1994).]
143. J. Hastbacka et al., *Cell* 78, 1073 (1994).
144. N. Risch, *Am. J. Hum. Genet.* 53 (suppl.), abstract 185 (1993); L. Kruglyak and E. S. Lander, submitted.
145. B. Karam et al., *Science* 245, 1073 (1989); J. Hastbacka et al., *Nat. Genet.* 2, 204 (1992).
146. In a population founded  $N$  generations ago, linkage disequilibrium should be detectable over distances

on the order of (100/M) cM. By studying  $k$  affected chromosomes, one should be able to localize a disease gene to a region on the order of (100/kM) cM. (Exact numbers would depend on the precise definition of detection of linkage disequilibrium and localization of a disease gene, but these estimates reflect the scaling with population age and number of affected chromosomes.)

147. Traditional construction of congenic strains by repeated backcrossing relies on the fact that an average of 50% of the undesired genome is lost at each generation. By using a complete genetic linkage map, however, one can identify those backcross progeny that have fortuitously lost a larger proportion of the undesired genome and breed them to create the next generation. In only

three to four generations, it is possible to eliminate essentially all of the undesired genome. For example, this has been performed to construct congenic strains for the Mm-1 region of mouse chromosome 4 (A. Moser and W. F. Dietrich, personal communication).

148. Complementation tests can be performed only between two alleles causing the same recessive phenotype. Accordingly, knockout experiments should target an allele  $A_1$  that causes a dominant (or partially dominant) phenotype when placed in trans to a second allele  $A_2$ ; the knockout allele would then be expected to fail to yield the dominant phenotype in the complementation test. Because current gene knockout protocols are limited to a few mouse strains such as 129, one

may first need to construct a congenic carrying the desired allele in such a strain before one can construct the appropriate knockout.

149. For an insightful historical account of the criticisms of Mendelian theory on the grounds that it cannot explain variation observed in nature, see W. B. Provine, *The Origins of Theoretical Population Genetics* (Univ. of Chicago Press, Chicago, 1971).

150. We thank L. Kruglyak and D. Siegmund for assistance concerning thresholds for significance and C. Amos, M. Boehnke, A. Chakravarti, F. Collins, R. Elston, W. Frankel, D. Finkel, S. Ghosh, S.-W. Guo, H. Jacob, J. Ott, A. Weder, A. Lynn, and members of the Lander laboratory for helpful comments on the manuscript. This work was supported in part by a grant from NIH (MG00098 to E.S.L.).

## AAAS-Newcomb Cleveland Prize

### To Be Awarded for a Report, Research Article, or an Article Published in *Science*

The AAAS-Newcomb Cleveland Prize is awarded to the author of an outstanding paper published in *Science*. The value of the prize is \$5000; the winner also receives a bronze medal. The current competition period began with the 3 June 1994 issue and ends with the issue of 26 May 1995.

Reports, Research Articles, and Articles that include original research data, theories, or syntheses and are fundamental contributions to basic knowledge or technical achievements of far-reaching consequence are eligible for consideration for the prize. The paper must be a first-time publication of the author's own work. Reference to pertinent earlier work by the author may be included to give perspective.

Throughout the competition period, readers are

invited to nominate papers appearing in the Reports, Research Articles, or Articles sections. Nominations must be typed, and the following information provided: the title of the paper, issue in which it was published, author's name, and a brief statement of justification for nomination. Nominations should be submitted to the AAAS-Newcomb Cleveland Prize, AAAS, Room 924, 1333 H Street, NW, Washington, DC 20005, and must be received on or before 30 June 1995. Final selection will rest with a panel of distinguished scientists appointed by the editor-in-chief of *Science*.

The award will be presented at the 1996 AAAS annual meeting. In cases of multiple authorship, the prize will be divided equally between or among the authors.